华章数学译丛

59

Pearson

Linear Algebra and Its Applications

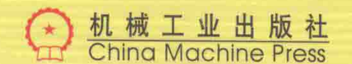
Fourth Edition

线性代数及其应用

（原书第4版）

[美] 戴维 C. 雷（David C. Lay）著

刘深泉 张万芹 陈玉珍 包东娥 陆博 译



华章数学译丛

59

Linear Algebra and Its Applications

Fourth Edition

线性代数及其应用

（原书第4版）

[美] 戴维 C. 雷 (David C. Lay) 著

刘深泉 张万芹 陈玉珍 包东娥 陆博 译

机械工业出版社
China Machine Press

图书在版编目（CIP）数据

线性代数及其应用（原书第4版）/（美）戴维 C. 雷（David C. Lay）著；刘深泉等译.
—北京：机械工业出版社，2017.3
（华章数学译丛）
书名原文：Linear Algebra and Its Applications, Fourth Edition

ISBN 978-7-111-55978-8

Ⅰ. 线… Ⅱ. ①戴… ②刘… Ⅲ. 线性代数-高等学校-教材 Ⅳ. O151.2

中国版本图书馆 CIP 数据核字（2017）第 029114 号

本书版权登记号：图字：01-2011-2303

线性代数是处理矩阵和向量空间的数学分支，在现代科学的各个领域都有应用．本书是一本优秀的现代教材，给出最新的线性代数基本介绍和一些有趣应用，目的是帮助学生掌握线性代数的基本概念及应用技巧，为后续课程的学习和工作实践奠定基础．主要内容包括线性方程组、矩阵代数、行列式、向量空间、特征值与特征向量、正交性和最小二乘法、对称矩阵和二次型、向量空间的几何学等．此外，本书包含大量的练习题、习题、例题等，便于读者参考．

本书内容深入浅出，论述清晰，适合作为高等院校理工科线性代数课程的教材，还可作为相关研究人员的参考书．

出版发行：机械工业出版社（北京市西城区百万庄大街 22 号　邮政编码：100037）
责任编辑：王春华　　责任校对：董纪丽
印　　刷：三河市宏图印务有限公司　　版　　次：2017 年 3 月第 1 版第 1 次印刷
开　　本：186mm×240mm　1/16　　印　　张：35.75
书　　号：ISBN 978-7-111-55978-8　　定　　价：79.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换
客服热线：（010）88378991　88361066　　投稿热线：（010）88379604
购书热线：（010）68326294　88379649　68995259　　读者信箱：hzjsj@hzbook.com

译 者 序

线性代数的应用范围覆盖自然科学和社会科学的各个方面，在处理高维问题时，向量空间和矩阵运算更是这些学科的理论基础和基本工具；在计算机和经济管理等学科，线性代数的重要性远远超过微积分. 在学习传统线性代数课程时，大学生经常遇到学习抽象理论和应用矩阵方法的矛盾；在今天深化改革开放的环境下，大学生更应该了解线性代数课程的世界发展趋势和最新教学进展.

David C. Lay 的教材《Linear Algebra and Its Applications》开创了学习和应用线性代数的新局面. 首先教材面向世界著名大学的优秀本科生，内容以矩阵运算和线性空间为主线，覆盖线性代数的基本理论和典型方法，创新习题要求借助数学软件完成. 其次教材内容满足不同学科应用线性代数的需要，理论深度超过国内线性代数教材的几乎所有内容. 不仅读者可以整体了解线性代数的知识体系和最新应用，而且各章节内容相对独立，教材学习可以根据学时和需要自由选择内容.

该教材自 2005 年中文翻译出版以来，深受读者欢迎. 很多教师将中文版作为线性代数的本科教材，也有读者作为学习线性代数的参考书. 目前原版英文教材已经多次修订，这次由华南理工大学的刘深泉教授，河南科技学院的张万芹、陈玉珍、包东娥、陆博老师完成新版本翻译. 由于译者水平的限制，书中难免存在不足之处，敬请各位同行批评指正，期望和读者一起，共同推动线性代数课程的向前发展.

刘深泉

2017 年 3 月 6 日

关 于 作 者

David C. Lay 在奥罗拉大学（伊利诺伊州）获得学士学位，在加利福尼亚大学（洛杉矶）获得硕士和博士学位. 自 1966 年以来，Lay 一直从事数学研究和数学教育工作，大部分时间是在马里兰大学帕克学院工作. 他还是阿姆斯特丹大学、阿姆斯特丹自由大学和德国凯泽斯劳滕大学的访问教授. 在泛函分析和线性代数方面，他已经发表的论文超过 30 篇.

作为美国 NSF 资助项目“线性代数课程研究小组”的核心成员，目前，Lay 是线性代数课程现代化的领导人. 此外，他还是几本数学教材的合著者，包括与 Angus E. Taylor 合著的《Introduction to Functional Analysis》，与 L. J. Goldstein 和 D. I. Schneider 合著的《Calculus and Its Applications》，以及与 D.Carlson，C.R.Johnson 和 A.D.Porter 合著的《Linear Algebra Gems—Assets for Undergraduate Mathematics》.

作为顶尖的教育家，Lay 教授获得过四所大学的杰出教学奖，包括 1996 年获得马里兰大学著名学者-教师称号. 1994 年，他获得美国数学联盟授予的著名大学数学教学奖. 他被大学生选为 Alpha-Lambda-Delta 国家荣誉专家协会和国家金钥匙荣誉协会的成员. 1989 年，奥罗拉大学授予他杰出校友荣誉. Lay 是美国数学会、加拿大数学会、国际线性代数协会、美国数学联盟、Sigma Xi 以及美国工业和应用数学学会成员.自从 1992 年以来，他成为多届数学科学基督联盟全国委员会成员.

前　言

学生和教师对本书前三版本的反响十分令人满意. 第 4 版在第 3 版的基础上为课程教学和软件技术应用提供了更多支持. 像以前一样，本书给出最新的线性代数基本介绍和一些有趣应用，使得已完成大学第二学期数学课程（如学完微积分）的学生容易接受.

本书的主要目的是帮助学生掌握以后课程学习所需要的基本概念和基本技能，教材的选题是根据“线性代数课程研究小组”的建议，该建议基于认真分析学生的实际需要和许多不同专业使用线性代数知识的共同点而提出. 希望这门课能够成为对大学生最有用和最有趣的数学课程之一.

新增内容

本版的主要目的是修订习题和拓展应用范围，这里既包括书中内容也包括网上资源。

1）超过 25%的习题是新增或修改的，尤其是计算习题. 习题的设置仍然是这本书最具特色的内容之一，并且新习题保持着与前三版一样的高标准. 精选的习题用于复习每个章节的所学内容，激发学生的学习兴趣，开拓新的思路，培养他们的自信心.

2）每章开头的 25%是新内容. 这部分给出一个线性代数应用的简短介绍，由此引出数学理论的发展. 在该章结束的部分又回到开始提到的应用.

3）新增的章：第 8 章（向量空间的几何学），提供了一个学生很感兴趣的新主题. 8.1 节、8.2 节和 8.3 节介绍基础几何学工具. 8.6 节用这些思想和方法研究贝塞尔曲线和曲面，并把它们用于工程和在线计算机绘图（Adobe Illustrator 和 Macromedia FreeHand）上，这四节内容大约需要 4 或 5 个课时（每课时 50 分钟）.

通常，本书的后续章节都会先对前一章节的关键思想、要点内容进行复习. 如果第 8 章部分视为前一章节，那么后继章节就会简要回顾 8.1～8.3 节的内容，然后重点介绍 8.4 节和 8.5 节的几何学.

4）学习指导是本书不可或缺的一部分，这部分都已作了修订，包括新的第 8 章. 同以前的版本一样，学习指导包含三分之一奇数号习题的全部详细解答，并对书中仅给出提示的奇数号习题也给出了解.

鲜明的特色

提前介绍重要概念

本书前 7 章介绍了许多建立在$\mathbb{R}^n$上的线性代数基本概念，然后从不同的观点逐步深入讨论. 接下来，用第 1 章给出的熟悉思想的自然扩展来泛化这些概念. 我认为，本教材的主要特色是全

书的难度一样.

矩阵乘法的现代观点

好的记号是关键，且教材反映科学家和工程师实际应用线性代数的方式. 本书在定义和证明中处理的是矩阵的列，而不是矩阵的元素，核心课题是将矩阵向量乘积 $\boldsymbol{Ax}$ 作为关于 $\boldsymbol{A}$ 的列的一个线性组合. 这种现代方法简化了许多论述，且将向量空间思想和线性系统的研究联系在一起.

线性变换

用线性变换作为线索贯穿整本教材，这增强了本书的几何趣味. 例如，在第 1 章，线性变换给出一个动态的和几何观点下的矩阵向量乘法.

特征值和动力系统

特征值的概念出现在第 5 章和第 7 章. 由于这一内容分散在数周的教学中，学生会比平常更容易吸收和复习这些关键概念. 特征值来源并应用于离散动力系统和连续动力系统，相关内容出现在 1.10 节、4.8 节、4.9 节和第 5 章的五节中. 在授课时可以选择不讲授第 4 章，而是在讲完 2.8 节和 2.9 节的内容以后直接进入第 5 章的学习. 这两节可选的学习内容给出了第 4 章中出现的向量空间的概念，为第 5 章的学习奠定了基础.

正交性和最小二乘法

与普通入门教材相比，本书对这些主题的讨论更全面. “线性代数课程研究小组”强调需要正交性和最小二乘问题的内容，这是由于正交性在计算机计算和线性代数的数值计算中起着重要作用，且实际工作中经常会出现不相容的线性方程组.

教学的特色

应用

广泛选取的应用说明了线性代数的作用，可以用于工程学、计算机科学、数学、物理学、生物学、经济学和统计学中解释基本原理和简化计算. 一些应用出现在单独的章节中；其他的应用是作为例题和习题而引入的. 此外，每一章的开头给出一个线性代数应用的简短介绍，由此引出数学理论的发展. 然后，在该章结束的部分又回到开始提到的应用.

重点强调几何特点

由于许多学生更容易接受形象化的概念，所以对书中的每个主要概念都给出几何解释. 本书包含较多的几何图形，且一些图形是以前的线性代数教材中没有出现过的.

例题

与大多数线性代数教材相比，本书有更多的例题，比平常课堂上更多. 由于例题清晰，步

骤详细，因此学生可以自学.

定理和证明

重要的结果以定理的形式给出，其他有用的事实放在方框中，便于参考. 大多数定理有正式证明，写法易于理解. 在少数情形中，仔细选取的例题证明中展示了基本计算过程. 一些常规的验证保留在习题中，这对学生是有益的.

练习题

在习题之前有几个仔细选取的练习题，其解答在习题之后给出. 这些练习题或者集中于习题中的潜在难点，或者给出做习题前的“热身”，且解答常包含有用的提示.

习题

提供的大量习题包含平常的计算题和需要深入思考的概念题，一些习题针对多年来我在学生作业中发现的概念难点. 每一个习题都按照课本中内容的顺序仔细排列；这样当每节的一部分内容讲授之后，就可以安排家庭作业. 习题的一个显著特色是数值计算不复杂，问题很快被“展开”，学生在数值计算上花费时间很少，习题主要是为了让学生理解教学内容而不是进行机械计算.

标有符号[M]的习题说明该题需要借助“矩阵软件”（计算机软件，如 MATLAB、Maple、Mathematica、MathCad 或者 Derive，或者有矩阵功能的可编程计算器，如那些由德州仪器制造的仪器）完成.

判断题

为鼓励学生阅读全部课本内容且深入思考，本书设计了 300 个简单的判断题，出现在 33 节内容之中，并放在计算题之后. 可以通过阅读课本内容来回答这些问题，从而使学生可以准备好回答随后的概念题. 学生在习惯了仔细阅读课本内容之后，会喜欢这类题目. 基于课堂测验以及与学生进行探讨，我决定不将答案放在课本中.（学习指导将指出在哪里有奇数号习题的答案.）补充的 150 道判断题（大部分在每章末尾）用于检验学生对内容的理解程度. 对大部分这类问题，教材中提供了简单的正确/错误回答，但是省略了答案的验证（通常需要进一步思考才可完成）.

写作题

写出严谨的数学论述，不仅对那些希望成为数学系研究生的学生，而且对所有学习线性代数的学生都十分必要. 本书包含的证明大多是习题答案的一部分. 需要简短证明的概念题，常包含可以帮助学生开始解题的提示. 对所有奇数的写作题，或者在课本的后面给出一个解答或提示，或者在后面描述的“学习指导”（Study Guide）中给出一个解答.

计算主题

本书强调计算机对科学和工程中线性代数的发展和实践的影响，书中有许多“数值计算的

注解”指出数值计算中出现的问题，以及理论概念（如矩阵求逆）和计算机实现（如 LU 分解）之间的区别.

网上支持

网站 www.pearsonhighered.com/lay 包含了学习本课程所需要的支持材料. 对学生而言，网站提供的复习资料和练习试卷（附答案）能帮助他们有效地复习书中的要点. 它们直接来自于近年来我所教授的课程. 每份复习资料给出了教材指定部分中重要的定义、定理和计算技巧.

各章应用

网站包含的 7 个案例研究扩展了每章开始所介绍的主题，增加了现实世界的数据，提供了进一步探索的机会. 另外，20 多个应用项目或者扩展了教材中的主题，或者介绍了新的应用，例如立方体样条、交通流量、飞机航线、运动比赛中的优势矩阵以及纠错码. 一些新的数学应用是积分方法、多项式根的求解、圆锥曲线、二次曲面和二元函数的极值. 线性代数的数值计算主题，如条件数、矩阵因式分解和求特征值的 QR 方法也包含在内. 在每个项目中都编有涉及大数据集的习题（因此需要使用计算机软件来求解）.

软件入门

如果课程中需要使用 MATLAB、Maple、Mathematica 或者 TI 计算器，可以参考网站上对软件使用的介绍. 另外，“学习指导”中也为第一次接触上述软件的用户提供了详细说明.

数据文件

对教材中的 900 道数值计算题、案例研究和应用项目，网站上提供了数百个相应的数据文件. 这些数据以多种格式存储，分别用于 MATLAB、Maple、Mathematica 以及 TI-83+/86/89 图形计算器. 对一道特定的题目，学生只需通过少量的键盘操作就可以得到矩阵和向量，从而减少了输入数据的错误，并可节省做作业的时间.

MATLAB 项目

这些探究性项目引导学生去发现线性代数中基础的数学与数值问题. 它们是由 Rink Smith 编写，基于佛罗里达大学计算线性代数课程的教学. 大约一半的项目探究的是有关基本概念，如列空间、对角化和正交投影；一些项目重点讨论数值计算问题，比如浮点计算、迭代法和 SVD；还有一些项目探究应用，如拉格朗日插值与马尔可夫链.

补充内容

学习指导

平装本的“学习指导”低价即可得到，我写的这个指导也作为该课程不可或缺的一部分. 它

给出每第三个奇数习题的详细解答，用以帮助学生检查作业. 它还为课本答案只包含“提示”的每一个奇数写作题提供了完整答案. 多次出现的“警告”用于提醒学生易犯的错误以及如何避免出错. MATLAB 框中介绍了一些需要使用的命令. “学习指导”的附录内包含 Maple、Mathematica 和 TI 图形计算器的对照信息（ISBN：0-321-38883-6）.

教师版

为方便教师，这个特殊版本包含所有习题的简单答案，课本开始在“给教师的注释”中给出了有关课程内容设计和组织的解释，以帮助教师安排课程. 该版本还提供了专为教师准备的其他支持材料（ISBN：0-321-38518-7）.

致谢

我真诚地感谢多年来以各种方式帮助我的许多人.

感谢 Israel Gohberg 和 Robert Ellis 长达 15 年的线性代数合作研究，他们帮助我形成了线性代数的观点. 与 David Carlson、Charles Johnson 和 Duane Porter 共同工作在“线性代数课程研究小组”是我的荣幸，在几个重要方面，他们关于线性代数教学的思想影响着本教材.

我诚挚地感谢以下审稿人，感谢他们细致入微的分析和创新性建议：田纳西理工大学的 Rafal Ablamowicz，西北大学的 John Alongi，华盛顿大学圣路易斯分校的 Brian E. Blank，佛罗里达州立大学的 Steven Bellenot，西蒙弗雷泽大学的 Vahid Dabbaghian-Abdoly，德雷塞尔大学的 Herman Gollwitzer，伍斯特大学的 James L. Hartman，德州大学奥斯丁分校的 David R. Kincaid，圣荷西州立大学的 Richard P. Kubelka，南卡罗莱纳州大学的 Douglas B. Meade，康涅狄格大学的 Martin Nikolov，佛罗里达大学的 Tim Olson，威廉玛莉学院的 Ilya M. Spitkovsky，杜兰大学的 Albert L. Vitter III. 该书是第 4 版，我要感谢在李伊大学的兄弟 Steven Lay，感谢他一直以来的帮助和鼓励，感谢他新修订了第 8 章. 还要感谢温思罗普大学的 Thomas Polaski. 关于每章介绍性实例的好的建议和帮助，我要感谢韦斯特蒙特学院的 Raymond Rosentrater. 另一位天才教授是华盛顿州立大学的 Judith McDonald，她开发了本书的很多新习题，对本书的帮助和热情是鼓舞人心的.

感谢提供技术支持的专家，是他们不辞劳苦地为第 4 版准备数据、给教师编写笔记、在“学习指导”中为学生编写技术笔记、将他们的项目和大家一起分享，他们是：泰勒大学的 Jeremy Case（MATLAB），南卡罗来纳大学的 Douglas Meade（Maple），西方浸信会学院的 Michael Miller（TI 计算器），以及卡罗尔学院的 Marie Vanisko（Mathematica）.

感谢专家 John Risley 与研究生 David Aulicino、Sean Burke 和 Hersh Goldberg，感谢他们对本书在线作业开发提供的专业技术帮助. 感谢对在线作业进行课堂测试的下列人员：澳大利亚国立大学的 Agnes Boskovitz、Malcolm Brooks、Elizabeth Ormerod、Alexander Isaev 和 John Urbas，马里兰大学蒙哥马利学院的 John Scott 和 Leben Wee，以及纽约州立大学布法罗分校的 Xingru

Zhang.

感激由 Blaise DeSesa、Jean Horn、Roger Lipsett、Paul Lorczak、Thomas Polaski、Sarah Streett 和 Marie Vanisko 提供的数学支持，他们对本书进行了仔细的校对.

最后，我由衷地感谢 Addison-Wesley 的工作人员在第 4 版的制作中提供的帮助：组稿编辑 Caroline Celano、高级内容编辑 Chere Bemelmans、副总编 Tamela Ambush、资深媒体编辑 Carl Cottrell、销售执行经理 Jeff Weidenaar、销售助理 Kendra Bassi、文本设计 Andrea Nix. 最后是三位好友——出版商 Gerg Tobin、前主编 Laurie Rosatone 和现主编 William Hoffman，他们从一开始就给予了明智的建议和鼓励，帮助解决了出版过程中的每一个问题，非常感谢他们！

David C. Lay

给学生的注释

这门课程是最有趣、最有价值的大学数学课程. 事实上，一些学生在毕业以后告诉我他们在大公司的工作中或工程研究生院的学习中还使用本教材作为参考书. 下面的注释给出一些建议和信息，有助于你掌握课本内容并且从中得到乐趣.

在线性代数中，概念和计算同样重要. 每个习题集开始的简单数值练习仅仅帮助你检查对基本步骤的理解. 以后，虽然计算机会进行数值计算，但你必须选取计算方法，知道如何解释结果，并且向其他人解释结果. 由于这个原因，课本中的许多习题要求你解释或验证计算. 书面解释经常是习题答案的一部分，对奇数习题，会有一个期望的解释或者好的提示. 在独立写出答案之前应尽量避免查阅这些答案，否则，你会认为你已理解实际上并不懂的问题.

为掌握线性代数的概念，你必须仔细地反复阅读本书. 新的术语用黑体标示，有时写在定义框中. 书的最后给出一个术语表，便于参考. 重要的命题以定理的形式给出或放在方框中. 最好阅读一下前言中对本书结构的介绍，以对课程的框架有初步的理解.

实际上，线性代数是一种语言，必须用学习外语的方法每天学习这种语言. 理解每一节的内容并不容易，除非你已透彻地学习教材且全部做出该节前面的练习，跟上课程的进度会帮助你节约很多时间和解决很多困惑！

数值计算的注解

希望你阅读课本中的“数值计算的注解”，即使你现在没有在学习过程中使用计算机或图形计算器. 在实际生活中，线性代数的大多数应用涉及一定数值误差限制下的数值计算，即使误差相当小.“数值计算的注解”会指出你以后工作中使用线性代数的潜在困难，如果现在学习了这些注释，你以后就会容易记起这些内容.

如果你对“数值计算的注解”有兴趣，以后可以学习一门线性代数的数值计算课程. 由于对计算机处理能力的更高需求，计算机科学家和数学家需要给出更快、更可靠的线性代数的数值算法，电子工程师需要设计出更快和更小的计算机去运行这些算法. 这是一个令人激动的领域，线性代数的第一堂课有助于你做好准备.

学习指导

为帮助你成功学习这门课程，我写了一本配合本书的“学习指导”（www.pearsonstore.com，0-321-38883-6）. 它不仅有助于你学习线性代数，而且说明了如何学习数学. 它包含每第三个

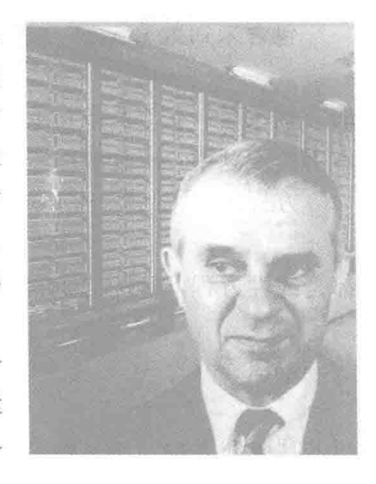
奇数习题的详细解答，以及课本答案仅有提示的奇数写作题的附加答案.“学习指导”还给出常见错误的警告、重点习题的有用提示和潜在考试题，且有一个单独的各章术语表（考试复习时很有用）. 此外，“学习指导”还介绍了如何使用 MATLAB、Maple、Mathematica 和 TI 图形计算器，使用这些技术可以节约完成作业的时间.

需要注意的是，之所以将“学习指导”与教材分开，是因为你必须在没有太多的帮助下独立完成作业. 多年的经验告诉我，太容易查到（如教材后面的）解答会妨碍学生数学能力的开发. 前几周的学习会培养出整个学期的学习习惯，并影响到学习效果. 请先阅读“如何学习线性代数”. 我的学生认为这些建议很有帮助，希望你也有同感.

目　录

第 1 章　线性代数中的线性方程组

介绍性实例　经济学与工程中的线性模型



1949 年夏末，哈佛大学教授列昂惕夫（Wassily Leontief）正在小心地将最后一部分穿孔卡片插入大学的 Mark Ⅱ 计算机. 这些卡片包含了美国经济的信息，包括了美国劳动统计局两年紧张工作所得到的总共 25 万多条信息. 列昂惕夫把美国经济分解为 500 个部门，例如煤炭工业、汽车工业、交通系统等等. 对每个部门，他写出了一个描述该部门的产出如何分配给其他经济部门的线性方程. 由于当时最大的计算机之一的 Mark Ⅱ 还不能处理所得到的包含 500 个未知数的 500 个方程的方程组，列昂惕夫只好把问题简化为包含 42 个未知数的 42 个方程的方程组.

为解列昂惕夫的 42 个方程，编写 Mark Ⅱ 计算机上的程序需要几个月的工作，他急于知道计算机解这个问题需要多长时间. Mark Ⅱ 计算机运算了 56 个小时，才得到最后的答案. 我们将在 1.6 节和 2.6 节中讨论这个解的性质.

列昂惕夫获得了 1973 年诺贝尔经济学奖，他打开了研究经济数学模型的新时代的大门. 1949 年在哈佛的工作标志着应用计算机分析大规模数学模型的开始. 从那以后，许多其他领域中的研究者应用计算机来分析数学模型. 由于所涉及的数据数量庞大，这些模型通常是线性的，即它们是用**线性方程组**描述的.

线性代数在应用中的重要性随着计算机功能的增大而迅速增加，而每一代新的硬件和软件引发了对计算机能力的更大需求. 因此，计算机科学就通过并行处理和大规模计算的爆炸性增长与线性代数密切联系在一起.

科学家和工程师正在研究大量极其复杂的问题，这在几十年前是不可想象的. 今天，线性代数对许多科学技术和工商领域中的学生的重要性可说超过了大学其他数学课程. 本书中的材料是在许多有趣领域中进一步研究的基础. 这里举出几个例子，以后将列举其他一些领域.

- *石油探测.* 当勘探船寻找海底石油储藏时，它的计算机每天要解几千个线性方程组. 方程组的地震数据从气喷枪的爆炸引起水下冲击波获得. 这些冲击波引起海底岩石的震动，并用几英里长的电缆拖在船后的地震测波器采集数据.
- *线性规划.* 许多重要的管理决策是在线性规划模型的基础上作出的，这些模型包含几百个变量，例如，航运业使用线性规划调度航班、监视飞机的飞行位置，或计划维修和机场运作.
- *电路.* 工程师使用仿真软件来设计电路和微芯片，它们包含数百万的晶体管. 这样的软件技术依赖于线性代数与线性方程组的方法.

>>>>>>>>>

线性方程组是线性代数的核心，本章使用它来引入线性代数的许多重要概念. 1.1 节和 1.2 节介绍求解线性方程组的一个系统方法，这个算法在全书的计算中都用到. 1.3 节和 1.4 节指出线性方程组等价于一个**向量方程**与**矩阵方程**. 这种等价性把向量的线性组合问题化为线性方程组的问题. 线性表示，线性无关和线性变换的基本概念将在本章后半部分研究，它们在整本书中起着关键的作用，并使我们体会到线性代数的魅力和威力.

1.1 线性方程组

包含未知数 $x_1,x_2,\cdots,x_n$ 的一个线性方程是形如

$$a_1x_1+a_2x_2+\cdots+a_nx_n=b \tag{1}$$

的方程，其中 b 与**系数** $a_1,a_2,\cdots,a_n$ 是实数或复数，通常是已知数. 下标 n 可以是任意正整数. 在本书的例题中，n 通常是在 2 与 5 之间. 在实际问题中，n 可以是 50，5000 或更大.

方程

$$4x_1-5x_2+2=x_1\text{和}x_2=2(\sqrt{6}-x_1)+x_3$$

都是线性方程，因为它们可以化为

$$3x_1-5x_2=-2\text{和}2x_1+x_2-x_3=2\sqrt{6}$$

方程

$$4x_1-5x_2=x_1x_2\text{和}x_2=2\sqrt{x_1}-6$$

都不是线性方程，因为第一个方程中包含 x_1x_2，第二个方程中包含 $\sqrt{x_1}$.

线性方程组是由一个或几个包含相同变量 $x_1,x_2,\cdots,x_n$ 的线性方程组成的. 例如，

$$\begin{aligned}2x_1-x_2+1.5x_3&=8\\x_1\qquad-4x_3&=-7\end{aligned} \tag{2}$$

线性方程组的一组**解**是一组数 $(s_1,s_2,\cdots,s_n)$，用这组数分别代替 $x_1,x_2,\cdots,x_n$ 时所有方程的两边相等. 例如，方程组（2）有一组解(5，6.5，3)，这是因为，在（2）中用这些值代替 x_1,x_2,x_3 时，方程组变成等式 8=8 和 $-7=-7$.

方程组所有可能的解的集合称为线性方程组的**解集**. 若两个线性方程组有相同的解集，则这两个线性方程组称为**等价**的. 也就是说，第一个方程组的每个解都是第二个方程组的解，第二方程组的每个解都是第一方程组的解.

求包含两个未知数的两个方程组成的方程组的解，等价于求两条直线的交点. 一个典型的例子是

$$\begin{aligned}x_1-2x_2&=-1\\-x_1+3x_2&=3\end{aligned}$$

这两个方程的图形都是直线，分别用 l_1 和 l_2 表示，数对 (x_1,x_2) 满足这两个方程当且仅当点 (x_1,x_2) 是这两条直线的交点. 容易验证，这个方程组有唯一的解(3，2)，如图 1-1 所示.

当然，两条直线不一定交于一个点，它们可能平行，也可能重合，重合的两条直线上的每个点都是交点. 图 1-2 是与下面两个方程组对应的图形：

a) $\begin{aligned} x_1-2x_2&=-1 \\ -x_1+2x_2&=\ \ 3 \end{aligned}$ b) $\begin{aligned} x_1-2x_2&=-1 \\ -x_1+2x_2&=\ \ 1 \end{aligned}$

图 1-1 和图 1-2 说明线性方程组的下列一般事实，这将在 1.2 节证明.

> 线性方程组的解有下列三种情况：
> 1. 无解.
> 2. 有唯一解.
> 3. 有无穷多解.

我们称一个线性方程组是**相容**的，若它有一个解或无穷多个解；称它是**不相容**的，若它无解.

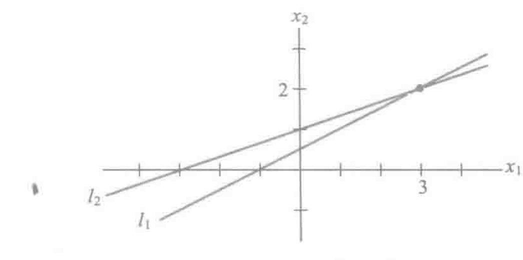


图 1-1 有唯一解

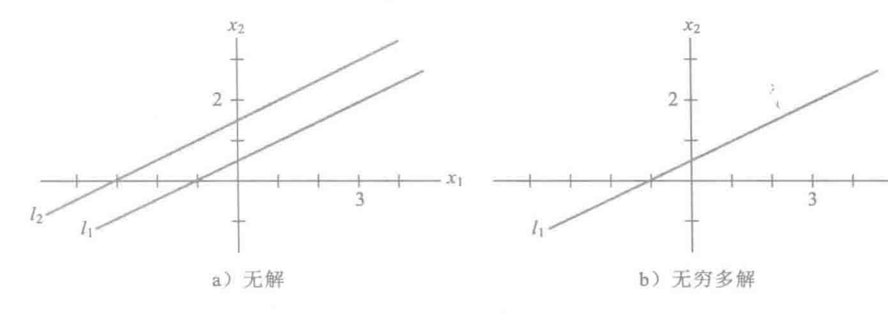


图 1-2

矩阵记号

一个线性方程组包含的主要信息可以用一个称为**矩阵**的紧凑的矩形阵列表示. 给出方程组

$$\begin{aligned} x_1-2x_2+\ x_3&=\ \ 0 \\ 2x_2-8x_3&=\ \ 8 \\ -4x_1+5x_2+9x_3&=-9 \end{aligned} \tag{3}$$

把每一个变量的系数写在对齐的一列中，矩阵

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix}$$

称为方程组（3）的**系数矩阵**，而

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix} \qquad (4)$$

称为它的**增广矩阵**.（第二行第一个元素为 0，因第二个方程可写成 $0 \cdot x_1 + 2x_2 - 8x_3 = 8.$）方程组的增广矩阵是把它的系数矩阵添上一列所得，这一列是由方程组右边常数组成的.

矩阵的**维数**说明它包含的行数和列数. 上面的增广矩阵（4）有 3 行 4 列，称为 3×4（读作 3 行 4 列）矩阵. 若 m, n 是正整数，一个 $m \times n$ **矩阵**是一个有 m 行 n 列的数的矩形阵列.（行数写在前面.）矩阵记号为解方程组带来方便.

线性方程组的解法

本节和下一部分给出了解线性方程组的一般方法. 基本的思路是把方程组用一个更容易解的等价方程组（即有相同解集的方程组）代替.

粗略地说，我们用方程组中第一个方程中含 x_1 的项消去其他方程中的含 x_1 的项. 然后用第二个方程中含 x_2 的项消去其他方程中含 x_2 的项，依此类推. 最后我们得到一个很简单的等价方程组.

用来化简线性方程组的三种基本变换：把某一个方程换成它与另一方程的倍数的和；交换两个方程的位置；把某一方程的所有的项乘以一个非零常数. 在例 1 之后，我们将说明经过这三种变换，为什么不改变方程组的解集.

例 1 解方程组（3）.

解 我们在消去未知数的同时用方程组与相应的矩阵形式表示出来以便比较.

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix}$$

保留第一个方程中的 x_1，把其他方程中的 x_1 消去. 为此，把第 1 个方程乘以 4，加到第 3 个方程上. 熟练之后可以通过心算完成：

$$\begin{array}{rr} 4\cdot[\text{方程1}]: & 4x_1 - 8x_2 + 4x_3 = 0 \\ +[\text{方程3}]: & -4x_1 + 5x_2 + 9x_3 = -9 \\ \hline [\text{新方程3}]: & -3x_2 + 13x_3 = -9 \end{array}$$

把原来的第三个方程用所得新方程代替：

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -3x_2 + 13x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

其次，把方程 2 乘以 1/2，使 x_2 的系数变成 1.

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ -3x_2 + 13x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

利用方程 2 中的 x_2 项消去方程 3 中的项 $-3x_2$，用心算计算如下：

$$\begin{array}{rr} 3\cdot[\text{方程2}]: & 3x_2-12x_3=12 \\ +[\text{方程3}]: & -3x_2+13x_3=-9 \\ \hline [\text{新方程3}]: & x_3=\ \ 3 \end{array}$$

所得的新方程组有三角形形状：[1]

$$\begin{aligned} x_1-2x_2+\ \ x_3&=0 \\ x_2-4x_3&=4 \\ x_3&=3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

现在我们想消去第一个方程中的项 $-2x_2$，不过先利用方程 3 消去第一个方程中的项 x_3 和第二个方程中的项 $-4x_3$ 更为有效. 这两个运算如下：

$$\begin{array}{rr} 4\cdot[\text{方程3}]: & 4x_3=12 \\ +[\text{方程2}]: & x_2-4x_3=\ \ 4 \\ \hline [\text{新方程2}]: & x_2\qquad\ \ =16 \end{array} \qquad \begin{array}{rr} -1\cdot[\text{方程3}]: & -x_3=-3 \\ +[\text{方程1}]: & x_1-2x_2+x_3=\ \ 0 \\ \hline [\text{新方程1}]: & x_1-2x_2\qquad =-3 \end{array}$$

这两次变换的结果如下：

$$\begin{aligned} x_1-2x_2\qquad &=-3 \\ x_2\qquad &=16 \\ x_3&=\ \ 3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

现在，在 x_3 的一列中只剩一项，我们回头来消去第一个方程中的 x_2 项. 把方程 2 的 2 倍加到方程 1，得到方程组：

$$\begin{cases} x_1 \qquad\qquad = 29 \\ \qquad x_2 \qquad = 16 \\ \qquad\qquad x_3 = 3 \end{cases} \qquad \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

我们已经得出结果：原方程组的唯一解是(29, 16, 3)，我们做了这么多计算，最好还是检验一下结果. 为证明(29, 16, 3)是方程组的解，把这些值代入原方程组的左边：

$$\begin{aligned} (29)-2(16)+\ \ (3)&=29-32+3=0 \\ 2(16)-8(3)&=32-24=8 \\ -4(29)+5(16)+9(3)&=-116+80+27=-9 \end{aligned}$$

结果与原方程组右边相同，所以(29, 16, 3)是原方程组的解（见图 1-3）.

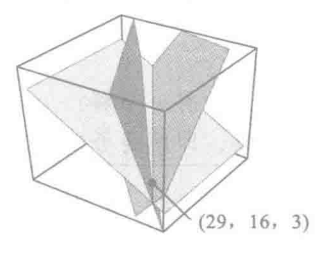


图 1-3　每个方程确定三维空间中的一个平面. 点(29, 16, 3)落在三个平面上 ■

[1] 直观的词“三角形”将在下节替代为更精确的术语.

例 1 说明了线性方程的变换对应于增广矩阵的行的变换. 前面所讲的三种基本变换对应于增广矩阵的下列变换.

行初等变换

1.（倍加变换）把某一行换成它本身与另一行的倍数的和.[1]
2.（对换变换）把两行对换.
3.（倍乘变换）把某一行的所有元素乘以同一个非零数.

行变换可施行于任何矩阵，不仅是对于线性方程组的增广矩阵. 我们称两个矩阵为**行等价**的，若其中一个矩阵可以经一系列行初等变换成为另一个矩阵.

重要的一点是行变换是可逆的. 若两行被对换，则再次对换它们就会还原为原来的状态. 若某一行乘以非零常数 c，则将所得的行乘以 $1/c$ 就得出原来的行. 最后，考虑涉及两行的倍加变换，例如第一行和第二行. 假设把第一行的 c 倍加到第二行得到新的第二行，那么“逆”变换就是把第一行的 $-c$ 倍加到（新的）第二行上就得到原来的第二行. 见本节末习题 29~32.

此时，我们更关注对一个线性方程组的增广矩阵进行行变换. 假设一个线性方程组经过行变换变成另一个新的方程组，考虑每一种行变换，容易看出，原方程组的任何一个解仍是新的方程组的一个解. 反之，因原方程组也可由新方程组经行变换得出，新方程组的每个解也是原方程组的解. 这就证明了下列事实.

若两个线性方程组的增广矩阵是行等价的，则它们具有相同的解集.

虽然例 1 看起来很长，经过一些练习你还是可以很快掌握的. 本书中习题的行变换通常是很简单的，因而你可以集中精力理解其中的思想和概念. 当然你必须熟练掌握这些变换，因为整本书都要用到它们.

本节的其他部分将介绍如何利用行变换来确定解集的情况，而无须完全求出解来.

存在与唯一性问题

在 1.2 节中我们将会知道，为什么一个线性方程组的解集可能不包含任何解、一个解或无穷多个解. 为确定某个方程组属于哪种情况，我们提出以下两个问题.

线性方程组的两个基本问题

1. 方程组是否相容，即它是否至少有一个解?
2. 若它有解，它是否只有一个解，即解是否唯一?

这两个问题将在整本书中以各种形式出现. 本节与下节中，我们将说明如何通过增广矩阵的行变换来回答这些问题.

---

[1] 通常把倍加变换说成：“把某一行的倍数加到另一行上”.

例 2 确定下列方程组是否有解

$$\begin{aligned} x_1-2x_2+\ \ x_3&=\ \ 0\\ 2x_2-8x_3&=\ \ 8\\ -4x_1+5x_2+9x_3&=-9 \end{aligned}$$

解 这就是例 1 中的方程组. 设我们已把方程组通过行变换变成三角形

$$\begin{aligned} x_1-2x_2+\ \ x_3&=0\\ x_2-4x_3&=4\\ x_3&=3 \end{aligned}\qquad \begin{bmatrix} 1 & -2 & 1 & 0\\ 0 & 1 & -4 & 4\\ 0 & 0 & 1 & 3 \end{bmatrix}$$

这时我们已经确定了 x_3，若把 x_3 的值代入方程 2，我们就会确定 x_2，因而可由方程 1 确定 x_1，所以解是存在的，即该方程组是相容的.（事实上，x_2 由方程 2 唯一确定，而 x_1 由方程 1 唯一确定，所以解是唯一的.） ■

例 3 确定下列方程组是否相容：

$$\begin{aligned} x_2-4x_3&=8\\ 2x_1-3x_2+2x_3&=1\\ 5x_1-8x_2+7x_3&=1 \end{aligned}\tag{5}$$

解 增广矩阵为：

$$\begin{bmatrix} 0 & 1 & -4 & 8\\ 2 & -3 & 2 & 1\\ 5 & -8 & 7 & 1 \end{bmatrix}$$

为使第一个方程包含 x_1 项，对换第一行与第二行：

$$\begin{bmatrix} 2 & -3 & 2 & 1\\ 0 & 1 & -4 & 8\\ 5 & -8 & 7 & 1 \end{bmatrix}$$

为消去第 3 个方程的项 $5x_1$，把第 1 行的$-5/2$ 倍加到第 3 行：

$$\begin{bmatrix} 2 & -3 & 2 & 1\\ 0 & 1 & -4 & 8\\ 0 & -1/2 & 2 & -3/2 \end{bmatrix}\tag{6}$$

其次，用第二个方程的 x_2 项消去第三个方程的 $-(1/2)x_2$ 项，把第 2 行的 1/2 倍加到第 3 行上：

$$\begin{bmatrix} 2 & -3 & 2 & 1\\ 0 & 1 & -4 & 8\\ 0 & 0 & 0 & 5/2 \end{bmatrix}\tag{7}$$

现在增广矩阵已成为三角形，我们来说明这个矩阵. 化为方程表示：

$$\begin{aligned} 2x_1-3x_2+2x_3&=1\\ x_2-4x_3&=8\\ 0&=5/2 \end{aligned}\tag{8}$$

方程 0=5/2 是 $0x_1+0x_2+0x_3=5/2$ 的简写. 这个三角形线性方程组显然是矛盾的，所以满足（8）

的未知数 x_1, x_2, x_3 的值是不可能存在的，因等式 0=5/2 不可能成立. 由于（8）和（5）有同样的解集，原方程组是不相容的（即无解），见图 1-4.

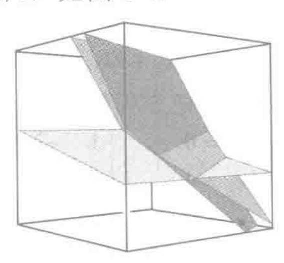


图 1-4　该方程组是不相容的，因为没有同时落在三个平面上的点 ■

注意（7）的增广矩阵，它的最后一行在三角形不相容方程组中是典型的.

数值计算的注解　在实际问题中线性方程组是通过计算机求解的. 对于方阵，计算机程序基本上是应用这里以及 1.2 节的消去法，稍做修正以改进精确度.

工商业中的大量线性代数问题运用浮点运算求解，数表示为小数形式：$\pm d_1 \cdots d_p \times 10^r$，$r$ 是整数，而小数点右面的数位为 8 至 16 位. 这种数的算术运算一般是有误差的. 其结果必须四舍五入（或舍去）为存储时需要的数位. “舍入误差”在输入像 1/3 这样的数时也会产生，因为它必须用近似的有限位小数表示. 幸运的是，浮点运算中的不精确性很少引起严重问题. 本书中关于数值计算的注解将会提醒你注意这些问题.

练习题

本书中练习题应该在做习题之前完成，它们的解答在每一节习题之后给出.

1. 用语言叙述解每个方程组时下一步应做的行变换.

a.
$$\begin{aligned} x_1 + 4x_2 - 2x_3 + 8x_4 &= 12 \\ x_2 - 7x_3 + 2x_4 &= -4 \\ 5x_3 - x_4 &= 7 \\ x_3 + 3x_4 &= -5 \end{aligned}$$

b.
$$\begin{aligned} x_1 - 3x_2 + 5x_3 - 2x_4 &= 0 \\ x_2 + 8x_3 \qquad &= -4 \\ 2x_3 \qquad &= 3 \\ x_4 &= 1 \end{aligned}$$

2. 某线性方程组的增广矩阵已经由行变换化为以下形式. 确定它是否有解.

$$\begin{bmatrix} 1 & 5 & 2 & -6 \\ 0 & 4 & -7 & 2 \\ 0 & 0 & 5 & 0 \end{bmatrix}$$

3. 数组(3，4，−2)是否为下列方程组的解：

$$\begin{aligned} 5x_1 - x_2 + 2x_3 &= 7 \\ -2x_1 + 6x_2 + 9x_3 &= 0 \\ -7x_1 + 5x_2 - 3x_3 &= -7 \end{aligned}$$

4. 当 h 和 k 取何值时下列方程组相容？

$$\begin{aligned} 2x_1 - x_2 &= h \\ -6x_1 + 3x_2 &= k \end{aligned}$$

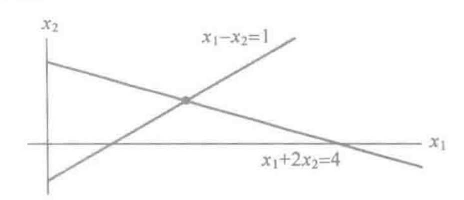
习题 1.1

利用对方程或增广矩阵的行初等变换解 1-4 题中的方程组. 依照本节中给出的消去过程.

1. $$\begin{aligned} x_1+5x_2&=7\\ -2x_1-7x_2&=-5 \end{aligned}$$

2. $$\begin{aligned} 3x_1+6x_2&=-3\\ 5x_1+7x_2&=10 \end{aligned}$$

3. 求在直线 $x_1+2x_2=4$和$x_1-x_2=1$ 上的点 (x_1,x_2)，见下图.



4. 求直线 $x_1+2x_2=-13$ 与 $3x_1-2x_2=1$ 的交点.

习题 5、6 中的矩阵是某个线性方程组的增广矩阵. 说明在解方程组时下一步应进行的行初等变换.

5. $$\begin{bmatrix} 1 & -4 & -3 & 0 & 7\\ 0 & 1 & 4 & 0 & 6\\ 0 & 0 & 1 & 0 & 2\\ 0 & 0 & 0 & 1 & -5 \end{bmatrix}$$

6. $$\begin{bmatrix} 1 & -6 & 4 & 0 & -1\\ 0 & 2 & -7 & 0 & 4\\ 0 & 0 & 1 & 2 & -3\\ 0 & 0 & 4 & 1 & 2 \end{bmatrix}$$

在 7~10 题中某个线性方程组的增广矩阵已由行变换化成如下所示，继续进行适当的行变换并说明原方程组的解集.

7. $$\begin{bmatrix} 1 & 7 & 3 & -4\\ 0 & 1 & -1 & 3\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & -2 \end{bmatrix}$$

8. $$\begin{bmatrix} 1 & -5 & 4 & 0 & 0\\ 0 & 1 & 0 & 1 & 0\\ 0 & 0 & 3 & 0 & 0\\ 0 & 0 & 0 & 2 & 0 \end{bmatrix}$$

9. $$\begin{bmatrix} 1 & -1 & 0 & 0 & -5\\ 0 & 1 & -2 & 0 & -7\\ 0 & 0 & 1 & -3 & 2\\ 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

10. $$\begin{bmatrix} 1 & 3 & 0 & -2 & -7\\ 0 & 1 & 0 & 3 & 6\\ 0 & 0 & 1 & 0 & 2\\ 0 & 0 & 0 & 1 & -2 \end{bmatrix}$$

解 11~14 题的方程组：

11. $$\begin{aligned} x_2+5x_3&=-4\\ x_1+4x_2+3x_3&=-2\\ 2x_1+7x_2+x_3&=-2 \end{aligned}$$

12. $$\begin{aligned} x_1-5x_2+4x_3&=-3\\ 2x_1-7x_2+3x_3&=-2\\ -2x_1+x_2+7x_3&=-1 \end{aligned}$$

13. $$\begin{aligned} x_1\qquad-3x_3&=8\\ 2x_1+2x_2+9x_3&=7\\ x_2+5x_3&=-2 \end{aligned}$$

14. $$\begin{aligned} 2x_1\qquad-6x_3&=-8\\ x_2+2x_3&=3\\ 3x_1+6x_2-2x_3&=-4 \end{aligned}$$

确定 15~16 题中的方程组是否相容，不必解出方程组.

15. $$\begin{aligned} x_1-6x_2\qquad\qquad&=5\\ x_2-4x_3+x_4&=0\\ -x_1+6x_2+x_3+5x_4&=3\\ -x_2+5x_3+4x_4&=0 \end{aligned}$$

16. $$\begin{aligned} 2x_1\qquad\qquad-4x_4&=-10\\ 3x_2+3x_3\qquad&=0\\ x_3+4x_4&=-1\\ -3x_1+2x_2+3x_3+x_4&=5 \end{aligned}$$

17. 三条直线 $2x_1+3x_2=-1$, $6x_1+5x_2=0$和$2x_1-5x_2=7$ 是否有一个交点？请解释.

18. 三条直线 $2x_1+4x_2+4x_3=4$, $x_2-2x_3=-2$和$2x_1+3x_2=0$ 是否有至少一个交点？请解释.

在 19~22 题中，确定 h 的值使矩阵是某一个相容方程组的增广矩阵：

19. $$\begin{bmatrix} 1 & h & 4\\ 3 & 6 & 8 \end{bmatrix}$$

20. $\begin{bmatrix} 1 & h & -5 \\ 2 & -8 & 6 \end{bmatrix}$

21. $\begin{bmatrix} 1 & 4 & -2 \\ 3 & h & -6 \end{bmatrix}$

22. $\begin{bmatrix} -4 & 12 & h \\ 2 & -6 & -3 \end{bmatrix}$

在 23~24 题中，或是直接引用本节的重要命题，略为改述（但依然成立），或被错误修改. 判断每个命题的真或假并给出理由. （若判断为真，指出教材中相似命题的出处，或参考的定义、定理. 若判断为假，指出引用或用错的命题出处，或举例说明.）类似的真/假问题会在本教材的许多章节中出现.

23. a. 每个行初等变换为可逆的.
    b. 5×6 矩阵有 6 行.
    c. 包含 n 个变量 $x_1,x_2,\cdots,x_n$ 的线性方程组的解集是一组数 $(s_1,s_2,\cdots,s_n)$，当用这组数代替 $x_1,x_2,\cdots,x_n$ 时，方程组中每个方程成为恒等式.
    d. 线性方程组的两个基本问题涉及存在性和唯一性.

24. a. 两个矩阵是行等价的，若它们有相同的行数.
    b. 对增广矩阵的行初等变换不改变相关的线性方程组的解集.
    c. 两个线性方程组是等价的，若它们有相同的解集.
    d. 不相容线性方程组有一个或更多解.

25. 求出包含 g，h 和 k 的方程组，使以下矩阵是相容方程组的增广矩阵.

$$\begin{bmatrix} 1 & -4 & 7 & g \\ 0 & 3 & -5 & h \\ -2 & 5 & -9 & k \end{bmatrix}$$

26. 设下面的方程组对所有 f 和 g 的可能取值都是相容的，则系数 c 和 d 有何特性？给出理由.

$$2x_1 + 4x_2 = f$$
$$cx_1 + dx_2 = g$$

27. 设 a,b,c,d 为常数，a 不为零，下面的方程组对所有 f 和 g 的可能取值都是相容的，则系数 a,b,c,d 有何特性？给出理由.

$$ax_1 + bx_2 = f$$
$$cx_1 + dx_2 = g$$

28. 给解集为 $x_1=3$, $x_2=-2$, $x_3=-1$ 的线性方程组构造三个不同的增广矩阵.

在 29~32 题中，求出把第一个矩阵变为第二个矩阵的行初等变换，并求出把第二个矩阵变为第一个矩阵的逆行初等变换.

29. $\begin{bmatrix} 0 & -2 & 5 \\ 1 & 3 & -5 \\ 3 & -1 & 6 \end{bmatrix}, \begin{bmatrix} 3 & -1 & 6 \\ 1 & 3 & -5 \\ 0 & -2 & 5 \end{bmatrix}$

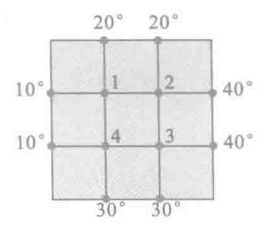
30. $\begin{bmatrix} 1 & 3 & -4 \\ 0 & -2 & 6 \\ 0 & -5 & 10 \end{bmatrix}, \begin{bmatrix} 1 & 3 & -4 \\ 0 & -2 & 6 \\ 0 & 1 & -2 \end{bmatrix}$

31. $\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 5 & -2 & 8 \\ 4 & -1 & 3 & -6 \end{bmatrix}, \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 5 & -2 & 8 \\ 0 & 7 & -1 & -6 \end{bmatrix}$

32. $\begin{bmatrix} 1 & 2 & -5 & 0 \\ 0 & 1 & -3 & -2 \\ 0 & 4 & -12 & 7 \end{bmatrix}, \begin{bmatrix} 1 & 2 & -5 & 0 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 0 & 15 \end{bmatrix}$

热传导的研究中，重要问题是确定某一块平板的稳态温度分布，假设已知边界上的温度分布，假设下图中的平板代表一条金属梁的截面，忽略垂直于该截面方向上的热传导，设 $T_1,T_2,\cdots,T_4$ 表示图中 4 个内部结点的温度. 某一节点的温度近似地等于 4 个与它最接近的结点（上、下、左、右）的平均值[㊀]，例如

$$T_1 = (10 + 20 + T_2 + T_4)/4 \text{ 或 } 4T_1 - T_2 - T_4 = 30$$



㊀ 参阅 *Heat and Mass Transfer*, F. M. White, Addison-Wesley Publishing, 1991, pp. 145-149.

33. 写出温度 $T_1, T_2, \cdots, T_4$ 所满足的方程组.

34. 解习题 33 中写出的方程组.（提示：为快速计算，在对换变换之前先对换第 1 行和第 4 行.）

练习题答案

1. a. 手工计算时，最好是把方程 3 与 4 对换，另一种办法是把方程 3 乘以 1/5，或把方程 4 化为它与方程 3 的−1/5 倍的和（不要利用方程 2 中的 x_2 项消去方程 1 中 $4x_2$，而要等到得到三角形矩阵且 x_3, x_4 项已从前 2 个方程消去之后）.

   b. 方程组已是三角形. 进一步的简化可利用第四个方程中的 x_4 项消去它上面的各个 x_4 项，现在适当的步骤是把方程 4 的 2 倍加到方程 1 上（此后再把方程 3 乘以 1/2，再消去其他的 x_3 项）.

2. 对应于该增广矩阵的方程组是

$$\begin{aligned} x_1 + 5x_2 + 2x_3 &= -6 \\ 4x_2 - 7x_3 &= 2 \\ 5x_3 &= 0 \end{aligned}$$

由第三个方程得到 $x_3 = 0$，这当然是允许的. 消去前两个方程中的 x_3 项，我们可以继续求出方程组的唯一解. 因此解是存在且唯一的. 注意比较本题与例 3.

3. 容易检验一组数是否是方程组的解. 把 $x_1 = 3$，$x_2 = 4$ 和 $x_3 = -2$ 代入方程组，求得

$$\begin{aligned} 5(3) - (4) + 2(-2) &= 15 - 4 - 4 = 7 \\ -2(3) + 6(4) + 9(-2) &= -6 + 24 - 18 = 0 \\ -7(3) + 5(4) - 3(-2) &= -21 + 20 + 6 = 5 \end{aligned}$$

虽然前两个方程满足了，但不满足第三个，所以 (3 , 4 , −2) 不是方程组的解. 注意代入时我们使用了括号. 建议你这么做是为了避免错误，见图 1-5.

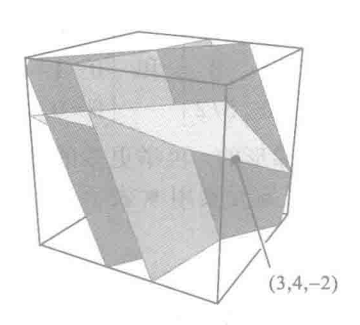


图 1-5　因为(3,4,−2)满足前两个方程，所以它落在前两个平面的交线上. 因为(3,4,−2)不满足所有三个方程，所以它没有落在三个平面上

4. 第二个方程代以它与第一个方程的 3 倍的和，方程组化为

$$\begin{aligned} 2x_1 - x_2 &= h \\ 0 &= k + 3h \end{aligned}$$

若 $k + 3h$ 非零，该方程组无解. 所以方程组对任意满足 $k + 3h = 0$ 的 h 和 k 的值有解.

1.2 行化简与阶梯形矩阵

本节我们将 1.1 节中的方法进一步精确化，变成行化简算法（也称行消去法），它可用来解任意线性方程组.[㊀] 而应用算法的第一部分，我们可以回答 1.1 节中提出的基本存在与唯一性问题.

这种算法可用于任意矩阵，不管它是否为某一方程组的增广矩阵. 所以本节的第一部分讨论任意矩阵. 首先我们引入两类重要的矩阵，它们包含 1.1 节中的“三角”矩阵，在以下的定义中，矩阵中非零行或列指矩阵中至少包含一个非零元素的行或列；非零行的**先导元素**是指该行中最左边的非零元素.

定义 一个矩阵称为**阶梯形**（或**行阶梯形**），若它有以下三个性质：

1. 每一非零行在每一零行之上.
2. 某一行的先导元素所在的列位于前一行先导元素的右面.
3. 某一先导元素所在列下方元素都是零.

若一个阶梯形矩阵还满足以下性质，称它为**简化阶梯形**（或**简化行阶梯形**）.

4. 每一非零行的先导元素是 1.
5. 每一先导元素 1 是该元素所在列的唯一非零元素.

若一个矩阵具有阶梯形（简化阶梯形），它就称为**阶梯形**（**简化阶梯形**）矩阵. 性质 2 说明先导元素构成阶梯形. 性质 3 其实是性质 2 的推论，不过我们把它列出来以示强调.

1.1 节中的“三角”矩阵，如

$$\begin{bmatrix}2 & -3 & 2 & 1\\0 & 1 & -4 & 8\\0 & 0 & 0 & 5/2\end{bmatrix} \text{ 和 } \begin{bmatrix}1 & 0 & 0 & 29\\0 & 1 & 0 & 16\\0 & 0 & 1 & 3\end{bmatrix}$$

都是阶梯形的，第二个矩阵是简化阶梯形的. 再举更多的例子.

例 1 下列矩阵都是阶梯形的，先导元素用 ■ 表示，它们可取任意的非零值，在*位置的元素可取任意值，包括零值.

$$\begin{bmatrix}■ & * & * & *\\0 & ■ & * & *\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{bmatrix} \quad \begin{bmatrix}0 & ■ & * & * & * & * & * & * & * & *\\0 & 0 & 0 & ■ & * & * & * & * & * & *\\0 & 0 & 0 & 0 & ■ & * & * & * & * & *\\0 & 0 & 0 & 0 & 0 & ■ & * & * & * & *\\0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & ■ & *\end{bmatrix}$$

■

下列矩阵是简化阶梯形的，因先导元素都是 1，且在每个先导元素 1 的上、下，各元素都是 0.

㊀ 我们的算法通常称为高斯消去法，类似的消去法大约在公元前 250 年由中国数学家应用. 这种方法在西方直到 19 世纪才由著名德国数学家 C.F.高斯发现. 德国工程师 W. 约当把它写在 1988 年的一本测地学著作中，这种方法才被普及.

$$\begin{bmatrix} 1 & 0 & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & * & 0 & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 1 & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \end{bmatrix}$$

■

一个矩阵可以**行化简**（即用行初等变换）变为阶梯形矩阵，但用不同的方法可化为不同的阶梯形矩阵. 然而，一个矩阵只能化为唯一的简化阶梯形矩阵. 下列定理将在书末附录 A 中给出证明.

定理 1 （简化阶梯形矩阵的唯一性）

每个矩阵行等价于唯一的简化阶梯形矩阵.

若矩阵 A 等价于阶梯形矩阵 U，称 U 为 A 的**阶梯形**（或行阶梯形）；若 U 是简化阶梯形，称 U 为 A 的**简化阶梯形**. 大部分矩阵程序用 RREF 作为简化（行）阶梯形的缩写，有些用 REF 作为（行）阶梯形的缩写.

主元位置

当矩阵经行变换化为阶梯形后，经进一步的行变换将矩阵化为简化阶梯形时，先导元素的位置并不改变. 因简化阶梯形是唯一的，当给定矩阵化为任何一个阶梯形时，先导元素总是在相同的位置上. 这些先导元素对应于简化阶梯形中的先导 1.

定义 *矩阵中的***主元位置***是 A 中对应于它的阶梯形中先导元素的位置.* **主元列***是 A 的含有主元位置的列.*

在例 1 中，符号（■）对应主元位置. 前四章中的许多基本概念都与矩阵中主元位置有联系.

例 2 把下列矩阵 A 用行变换化为阶梯形，确定主元列.

$$A = \begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix}$$

解 利用 1.1 节的方法. 最左边的非零列的第一个元素就是第一个主元位置. 这个位置必须放一个非零元，即主元素. 最好将第一行与第四行对换，这样可以避免分数运算.

$$\begin{bmatrix} \overset{\text{主元}}{1} & 4 & 5 & -9 & -7 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ \underset{\uparrow\,\text{主元列}}{0} & -3 & -6 & 4 & 9 \end{bmatrix}$$

把第一行的倍数加到其他各行，以使主元 1 下面各元素变成 0. 第二行的主元位置必须尽量靠

左，即在第二列. 我们选择这里的 2 作为第二个主元.

主元

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \quad (1)$$

新的主元列

把第二行的−5/2 倍加到第三行，3/2 倍加到第 4 行.

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 & 0 \end{bmatrix} \quad (2)$$

（2）中的矩阵与 1.1 节所遇到的不同，这里没有办法在第 3 列中找到先导元素！我们不能利用第一行或第二行，否则会破坏已产生的阶梯形的先导元素的排列. 然而若我们对换第 3 行和第 4 行，我们可在第 4 列产生先导元素.

主元

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{一般形式} \qquad \begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

主元列

此矩阵已是阶梯形. 第 1、2、4 列是主元列.

主元位置

$$A = \begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix} \quad (3)$$

主元列

■

如例 2 所示，**主元**就是在主元位置上的非零元素，用来通过行变换把下面的元素化为 0，例 2 中的主元是 1，2，−5，注意这些元素与矩阵 A 中同一位置的元素不相同，如（3）式所示.

根据例 2，我们给出一个有效的算法，变换矩阵成阶梯形或简化阶梯形. 认真掌握这一算法将使你获益匪浅.

行化简算法

下列算法包含四个步骤，它产生一个阶梯形矩阵，第五步产生简化阶梯形矩阵. 我们用一个实例来说明这一算法.

例 3 用行初等变换把下列矩阵先化为阶梯形，再化为简化阶梯形.

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

解

第一步，由最左的非零列开始. 这是一个主元列. 主元位置在该列顶端.

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

主元列（第 1 列）

第二步，在主元列中选取一个非零元作为主元. 若有必要的话，对换两行使这个元素移到主元位置上.

对换第 1、3 两行（也可对换 1、2 两行）

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

主元：第 1 行第 1 列的 3

第三步，用倍加行变换将主元下面的元素变成 0.

我们当然可以把第 1 行除以主元 3. 但这里第 1 列有两个 3，我们只需把第 1 行的–1 倍加到第 2 行.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

主元：第 1 行第 1 列的 3

第四步，暂时不管包含主元位置的行以及它上面的各行，对剩下的子矩阵使用上述的三个步骤直到没有非零行需要处理为止.

暂时不看第一行，第一步指出，第 2 列是下一个主元列，第二步，我们选择该列中“顶端”的元素作为主元.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

主元：第 2 行第 2 列的 2；新主元列：第 2 列

对第三步，我们可先把子矩阵的“顶行”除以主元 2. 不过也可以把这一行的–3/2 倍加到下面的一行. 这就得到

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

暂时不看第二个主元所在的行，我们剩下一个只有一行的新子矩阵.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

主元

新的子矩阵已不需要处理了，我们已得到整个矩阵的阶梯形. 若我们需要简化阶梯形，进行下一个步骤.

> 第五步，由最右面的主元开始，把每个主元上方的各元素变成 0. 若某个主元不是 1，用倍乘变换将它变成 1.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{matrix} \leftarrow 行1+(-6)\cdot 行3 \\ \leftarrow 行2+(-2)\cdot 行3 \\ \end{matrix}$$

下一个主元在第 2 行，将这行除以这个主元.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{matrix} \\ \leftarrow 行乘以1/2 \\ \end{matrix}$$

将第 2 行的 9 倍加到第 1 行

$$\begin{bmatrix} 3 & 0 & -6 & 9 & 0 & -72 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{matrix} \leftarrow 行1+9\cdot 行2 \\ \\ \end{matrix}$$

最后将第 1 行除以主元 3

$$\begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{matrix} \\ \leftarrow 行乘以1/3 \\ \end{matrix}$$

这就是原矩阵的简化阶梯形. ■

第 1~4 步称为行化简算法的**向前步骤**，产生唯一的简化阶梯形的第 5 步，称为**向后步骤**.

数值计算的注解 在第 2 步中，计算机程序通常选择一列中绝对值最大的元素作为主元. 这种方法通常称为**部分主元法**，可以减少计算中的舍入误差.

线性方程组的解

行化简算法应用于方程组的增广矩阵时，可以得出线性方程组解集的一种显式表示法.

例如，设某一个线性方程组的增广矩阵已经化为等价的简化阶梯形

$$\begin{bmatrix} 1 & 0 & -5 & 1 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

因为增广矩阵有 4 列，所以有 3 个未知数，对应的线性方程组是

$$\begin{aligned} x_1 \quad -5x_3 &= 1 \\ x_2 + \ x_3 &= 4 \\ 0 &= 0 \end{aligned} \qquad (4)$$

对应于主元列的变量 x_1 和 x_2 称为**基本变量**[㊀]. 其他变量如 x_3，称为**自由变量**.

当一个线性方程组是相容的，如方程组（4）解集可以用显式表示，只要把方程的简化形式解出来，用自由变量表示基本变量即可. 由于简化阶梯形使每个基本变量仅包含在一个方程中，这是很容易的. 在方程组（4）中，我们可由第 1 个方程解出 x_1，第 2 个方程解出 x_2（第 3 个方程对未知数没有任何限制，可以不管它）.

$$\begin{cases} x_1 = 1 + 5x_3 \\ x_2 = 4 - x_3 \\ x_3\text{是自由变量} \end{cases} \qquad (5)$$

我们说 x_3 是自由变量，是指它可取任意的值. 当 x_3 的值选定后，由（5）中的前 2 个方程就可以确定 x_1 和 x_2 的值，例如，当 $x_3 = 0$，得出解（1，4，0）；当 $x_3 = 1$，得出解（6，3，1），x_3 的不同选择确定了方程组的不同的解，方程组的每个解由 x_3 的值的选择来确定.

（5）式给出的解称为方程组的**通解**，因为它给出了所有解的显式表示.

例 4 求方程组的解，该方程组的增广矩阵已经化为

$$\begin{bmatrix} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

解 该矩阵已是阶梯形，但我们在解出基本变量前仍需把它化为简化阶梯形，记号“~”表示它前面和后面的两个矩阵是行等价的（译者注：该记号在中文教科书中并未通用）.

$$\begin{bmatrix} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 2 & -5 & 0 & 10 \\ 0 & 0 & 2 & -8 & 0 & 10 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 6 & 2 & -5 & 0 & 10 \\ 0 & 0 & 1 & -4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & -4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

增广矩阵有 6 列，所以原方程组有 5 个变量，对应的方程组为

㊀ 某些书称为先导变量，因为它们对应于包含先导元素的主元列.

$$\begin{aligned} x_1 + 6x_2 \quad + 3x_4 \quad &= 0 \\ x_3 - 4x_4 \quad &= 5 \\ x_5 &= 7 \end{aligned} \tag{6}$$

矩阵的主元列是第 1，3，5 列，基本变量为 x_1， x_3， x_5，剩下的变量 x_2 和 x_4 为自由变量，解出基本变量，我们得通解为

$$\begin{cases} x_1 = -6x_2 - 3x_4 \\ x_2\text{为自由变量} \\ x_3 = 5 + 4x_4 \\ x_4\text{为自由变量} \\ x_5 = 7 \end{cases} \tag{7}$$

注意，由方程组（6）的第 3 个方程，x_5 的值是确定的. ■

解集的参数表示

解集的表示式（5）和（7）称为解集的参数表示，其中自由变量作为参数. *解方程组*就是要求出解集的这种参数表示或确定它无解.

当一个方程组是相容的，且具有自由变量，则它的解集具有多种参数表示. 例如，在方程组（4）中，我们可以把方程 2 的 5 倍加到方程 1，得等价方程组

$$\begin{aligned} x_1 + 5x_2 \quad\quad &= 21 \\ x_2 + x_3 &= \ \ 4 \end{aligned}$$

这时可把 x_2 看作参数，用 x_2 表示 x_1 和 x_3，得到解集的第一种表示法. 不过，我们总是约定使用自由变量作为参数来表示解集（本书末尾的习题解答也采用这一约定）.

当方程组是不相容时，解集是空集，无论方程组是否有自由变量，此时，解集无参数表示.

回代

考虑下列方程组，它的增广矩阵已是阶梯形，但还不是简化阶梯形：

$$\begin{aligned} x_1 - 7x_2 + 2x_3 - 5x_4 + 8x_5 &= 10 \\ x_2 - 3x_3 + 3x_4 + \ \ x_5 &= -5 \\ x_4 - \ \ x_5 &= \ \ 4 \end{aligned}$$

计算机程序通常用回代法解此方程组，而不是求它的简化阶梯形. 也就是说，程序先解第 3 个方程，用 x_5 表示 x_4，并把此表达式代入第 2 个方程，从中解出 x_2，最后把 x_2 和 x_4 的表达式代入第 1 个方程解出 x_1.

我们的矩阵算法，即行化简算法的向后步骤，它求出简化阶梯形，与回代法所需算术运算次数相同. 但矩阵算法通常减少了手算时出错的可能性. 我强烈建议你仅使用简化阶梯形来解方程组！与本书配合的学习指导书给出一些好的建议帮助你更快更准确地解方程组.

数值计算的注解 *一般地，行化简算法的向前步骤比向后步骤需要更多运算. 解方程组的算法通常用浮算来衡量. 一个浮算（flop 或浮点运算）就是两个浮点实数进行一*

次算术运算（+，-，*，/）.[⊖] 对一个 $n\times(n+1)$ 矩阵，化简为阶梯形大约需要 $2n^3/3+n^2/2-7n/6$ 次浮算（当 n 相当大，比如说 $n\geqslant 30$ 时，大约是 $2n^3/3$ 次浮算），而进一步化为简化阶梯形大约最多只需 n^2 次浮算.

存在与唯一性问题

虽然非简化的阶梯形并不适于解线性方程组，这种形式对于回答 1.1 节中提出的两个基本问题已经足够了.

例 5 确定下列线性方程组的解是否存在且唯一

$$\begin{aligned} 3x_2-6x_3+6x_4+4x_5&=-5\\ 3x_1-7x_2+8x_3-5x_4+8x_5&=9\\ 3x_1-9x_2+12x_3-9x_4+6x_5&=15 \end{aligned}$$

解 该方程组的增广矩阵在例 3 中化简为

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15\\ 0 & 2 & -4 & 4 & 2 & -6\\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \tag{8}$$

基本变量是 x_1,x_2 和 x_5；自由变量是 x_3 和 x_4. 这里没有类似 0=1 的造成不相容方程组的方程，所以我们可用回代法求解. 但解的存在性在方程（8）中已经清楚了. 同时，解不是唯一的，因为有自由变量存在. x_3 和 x_4 的每一种选择都确定一组解，所以此方程组有无穷多组解. ■

当一个方程组化为阶梯形，且不包含形如 $0=b$ 的方程，其中 $b\neq 0$，每个方程包含一个基本变量，它的系数非零. 或者这些基本变量已完全确定（此时无自由变量），或者至少有一个基本变量可用一个或多个自由变量表示，对前一种情形，有唯一的解；对后一种情形，有无穷多个解（对应自由变量的每一个选择都有一个解）.

上述讨论证明了以下定理

定理 2 （存在与唯一性定理）

线性方程组相容的充要条件是增广矩阵的最右列不是主元列. 也就是说，增广矩阵的阶梯形没有形如

$$[0\ \cdots\ 0\ \ b]\quad b\neq 0$$

的行，若线性方程组相容，它的解集可能有两种情形：（ⅰ）当没有自由变量时，有唯一解；（ⅱ）若至少有一个自由变量，有无穷多解.

以下是求解线性方程组的步骤.

[⊖] 传统意义上，浮算仅表示一次乘法或除法，因为加法和减法需时很少，可以忽略，这里定义的浮算现在使用较多，这是由于计算机算术运算能力的提高. 见 Golub and Van Loan, *Matrix Computations*, 2nd ed. (Baltimore: The Johns Hopkins Press, 1989), pp. 19-20.

应用行化简算法解线性方程组

1. 写出方程组的增广矩阵.
2. 应用行化简算法把增广矩阵化为阶梯形. 确定方程组是否有解，如果没有解则停止；否则进行下一步.
3. 继续行化简算法得到它的简化阶梯形.
4. 写出由第3步所得矩阵所对应的方程组.
5. 把第4步所得的每个方程改写为用自由变量表示基本变量的形式.

练习题

1. 求出下列增广矩阵对应的方程组的解

$$\begin{bmatrix} 1 & -3 & -5 & 0 \\ 0 & 1 & 1 & 3 \end{bmatrix}$$

2. 求出下列方程组的通解

$$\begin{aligned} x_1 - 2x_2 - x_3 + 3x_4 &= 0 \\ -2x_1 + 4x_2 + 5x_3 - 5x_4 &= 3 \\ 3x_1 - 6x_2 - 6x_3 + 8x_4 &= 2 \end{aligned}$$

习题 1.2

在习题1~2中，确定哪些矩阵是简化阶梯形，哪些仅是阶梯形.

1. a. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$　b. $\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$　d. $\begin{bmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 2 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}$

2. a. $\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$　b. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

c. $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$　d. $\begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

行化简习题3~4中的矩阵为简化阶梯形. 在最终的矩阵和原始矩阵中圈出主元位置，指出主元列.

3. $\begin{bmatrix} 1 & 2 & 4 & 8 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 9 & 12 \end{bmatrix}$　4. $\begin{bmatrix} 1 & 2 & 4 & 5 \\ 2 & 4 & 5 & 4 \\ 4 & 5 & 4 & 2 \end{bmatrix}$

5. 给出一个非零 2×2 矩阵可能的阶梯形，用例1中的符号■，*和0表示.

6. 对一个非零 3×2 矩阵，重复习题5.

在习题7~14中，给出线性方程组的增广矩阵，求其通解.

7. $\begin{bmatrix} 1 & 3 & 4 & 7 \\ 3 & 9 & 7 & 6 \end{bmatrix}$　8. $\begin{bmatrix} 1 & -3 & 0 & -5 \\ -3 & 7 & 0 & 9 \end{bmatrix}$

9. $\begin{bmatrix} 0 & 1 & -2 & 3 \\ 1 & -3 & 4 & -6 \end{bmatrix}$　10. $\begin{bmatrix} 1 & -2 & -1 & 4 \\ -2 & 4 & -5 & 6 \end{bmatrix}$

11. $\begin{bmatrix} 3 & -2 & 4 & 0 \\ 9 & -6 & 12 & 0 \\ 6 & -4 & 8 & 0 \end{bmatrix}$

12. $\begin{bmatrix} 1 & 0 & -9 & 0 & 4 \\ 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

13. $\begin{bmatrix} 1 & -3 & 0 & -1 & 0 & -2 \\ 0 & 1 & 0 & 0 & -4 & 1 \\ 0 & 0 & 0 & 1 & 9 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

14. $\begin{bmatrix} 1 & 0 & -5 & 0 & -8 & 3 \\ 0 & 1 & 4 & -1 & 0 & 6 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

习题15~16使用例1中阶梯形矩阵的符号给出线性方程组的增广矩阵，判断每个矩阵对应的方程组是否相容. 如果方程组相容，判断解是否唯一.

15. a. $\begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 \end{bmatrix}$　b. $\begin{bmatrix} 0 & \blacksquare & * & * & * \\ 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & \blacksquare & 0 \end{bmatrix}$

16. a. $\begin{bmatrix} \blacksquare & * & * \\ 0 & \blacksquare & * \\ 0 & 0 & \blacksquare \end{bmatrix}$　b. $\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & \blacksquare & * \end{bmatrix}$

在习题17~18中，确定h的值，使所给矩阵是一个相容方程组的增广矩阵.

17. $\begin{bmatrix} 1 & -1 & 4 \\ -2 & 3 & h \end{bmatrix}$　18. $\begin{bmatrix} 1 & -3 & 1 \\ h & 6 & -2 \end{bmatrix}$

在习题19~20中，确定h和k的值，使方程组(a)无解，(b)有唯一解，(c)有多解，给出每种情况的答案.

19. $\begin{aligned} x_1 + hx_2 &= 2 \\ 4x_1 + 8x_2 &= k \end{aligned}$　20. $\begin{aligned} x_1 - 3x_2 &= 1 \\ 2x_1 + hx_2 &= k \end{aligned}$

在习题21~22中，判断每个命题的真假，给出理由.㊀

21. a. 某些矩阵只要用不同的行运算次序就可行化简为不止一个简化阶梯形矩阵.
    b. 行化简算法只能用于线性方程组的增广矩阵.
    c. 线性方程组中的基本变量是系数矩阵中主元列对应的变量.
    d. 求线性方程组的解集的参数表示就是解这个方程组.
    e. 若某个增广矩阵的一行是[0 0 0 5 0]，则对应的线性方程组是不相容的.
22. a. 矩阵的阶梯形是唯一的.
    b. 若增广矩阵的每一列都包含一个主元，则相应的方程组是相容的.
    c. 矩阵的主元位置依赖于在行化简算法中是否使用行对换变换.
    d. 线性方程组的通解是所有解的一个显式表示.
    e. 当方程组具有自由变量时，解集一定包含许多解.
23. 设包含4个未知数的4个方程的系数矩阵每列有一个主元. 说明为什么方程组有唯一解.
24. 设一个线性方程组的3×5增广矩阵的第5列是主元列. 该方程组是否相容，为什么？
25. 设一个线性方程组的系数矩阵每行有一个主元位置，说明为什么方程组是相容的.
26. 设一个方程组的3×5系数矩阵有3个主元列，这个方程组是否相容，为什么？
27. 利用主元列的概念重述定理2的最后一句:"若线性方程组是相容的，则解是唯一的当且仅当________."
28. 为了知道一个方程组是相容的且具有唯一解，你需要知道它的增广矩阵的主元列的什么情况？
29. 若线性方程组的方程个数少于未知数个数，它称为*欠定方程组*. 设一个欠定方程组是相容的，说明它为什么会有无穷多解.
30. 给出一个含有3个未知数和2个方程的不相容的欠定方程组的例子.
31. 若线性方程组的方程个数多于未知数个数，它称为*超定方程组*. 这样的方程组是否可能相容？用一个含2个未知数，3个方程的方程组说明你的答案.
32. 设一个$n\times(n+1)$矩阵用行化简算法化为简化阶梯形，当$n=20$和$n=200$时总的运算(浮算)次数中，向后步骤占了多少比例？

设实验数据用平面上一些点表示. 这些数据的一个**插值多项式**是其图像通过这些点的一个多项式. 在科学工作中，这样的多项式可用来估计在已知数据之间的一些数值. 另一个应用是在计算

㊀ 判断命题的真假在许多章节中出现，论证的方法在1.1节习题23~24之前的说明中已经指出.

机屏幕绘制造图像. 求这种插值多项式的一种方法是解线性方程组.

33. 求数据（1,6），（2,15），（3,28）的插值多项式 $p(t)=a_0+a_1t+a_2t^2$. 即求 a_0,a_1,a_2 使

$$a_0+a_1(1)+a_2(1)^2=6$$
$$a_0+a_1(2)+a_2(2)^2=15$$
$$a_0+a_1(3)+a_2(3)^2=28$$

34. [M]在一次风洞实验中，空气对飞机的阻力在不同速度下为

速度 (100ft/sec)	0	2	4	6	8	10
空气阻力 (100lb)	0	2.90	14.8	39.6	74.3	119

求这些数据的插值多项式，估计飞机速度为750ft/sec时的空气阻力用

$$p(t)=a_0+a_1t+a_2t^2+a_3t^3+a_4t^4+a_5t^5$$

若你用低于5次的多项式去计算，结果如何？（例如，试用3次多项式.）[⊖]

练习题答案

1. 增广矩阵的简化阶梯形和相应的方程组是

$$\begin{bmatrix}1&0&-2&9\\0&1&1&3\end{bmatrix} \quad 和 \quad \begin{aligned}x_1\quad -2x_3&=9\\x_2+\ x_3&=3\end{aligned}$$

基本变量是 x_1 与 x_2，通解为（见图1-6）：

$$\begin{cases}x_1=9+2x_3\\x_2=3-x_3\\x_3\text{是自由变量}\end{cases}$$

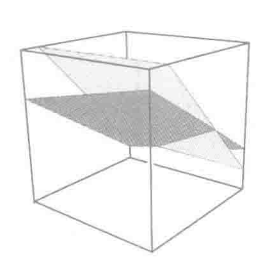


图1-6 方程组的通解是空间中两个平面相交的直线

注意：要点是一般解要描述每个变量，参数要明显标出. 下面的写法是不正确的：

$$\begin{cases}x_1=9+2x_3\\x_2=3-x_3\\x_3=3-x_2 \quad \text{不正确的解}\end{cases}$$

在这种写法中，似乎 x_2 和 x_3 都是自由变量，这当然是不对的.

2. 当我们行化简方程组的增广矩阵，得

$$\begin{bmatrix}1&-2&-1&3&0\\-2&4&5&-5&3\\3&-6&-6&8&2\end{bmatrix}\sim\begin{bmatrix}1&-2&-1&3&0\\0&0&3&1&3\\0&0&-3&-1&2\end{bmatrix}\sim\begin{bmatrix}1&-2&-1&3&0\\0&0&3&1&3\\0&0&0&0&5\end{bmatrix}$$

⊖ 标有记号[M]的习题是设计为利用“矩阵程序”来求解的，例如MATLAB、Maple、Mathemática、MathCad、Derive或有矩阵计算功能的计算器，如德州仪器公司或惠普公司所制造的.

所得阶梯矩阵说明方程是不相容的，因它的最右列是主元列；第 3 行相应于方程 0=5，因此不必再进行任何行变换. 注意，自由变量在此问题中并不起作用，因为方程组是不相容的.

1.3 向量方程

线性方程组的重要性质都可用向量概念与符号来描述. 本节把通常的线性方程组与向量方程联系起来. 名词“向量”出现在各种数学和物理教科书中，在第 4 章我们将讨论“向量空间”. 在此以前，我们用向量表示一组数. 这种简单的思想使得我们能够尽快地将它们应用于有趣和重要的问题.

$\mathbb{R}^2$ 中的向量

仅含一列的矩阵称为**列向量**，或简称**向量**，包含两个元素的向量如下所示.

$$\boldsymbol{u}=\begin{bmatrix}3\\-1\end{bmatrix},\ \boldsymbol{v}=\begin{bmatrix}0.2\\0.3\end{bmatrix},\ \boldsymbol{w}=\begin{bmatrix}w_1\\w_2\end{bmatrix}$$

这里 w_1 和 w_2 是任意实数. 所有两个元素的向量的集记为 $\mathbb{R}^2$，$\mathbb{R}$ 表示向量中的元素是实数，而指数 2 表示每个向量包含两个元素.[⊖]

$\mathbb{R}^2$ 中两个向量**相等**，当且仅当对应元素相等. 因此，$\begin{bmatrix}4\\7\end{bmatrix}$ 和 $\begin{bmatrix}7\\4\end{bmatrix}$ 是不相等的. 我们称 $\mathbb{R}^2$ 中向量是实数的有序对.

给定 $\mathbb{R}^2$ 中两个向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$，它们的和 $\boldsymbol{u}+\boldsymbol{v}$ 是把对应元素相加所得向量. 例如，

$$\begin{bmatrix}1\\-2\end{bmatrix}+\begin{bmatrix}2\\5\end{bmatrix}=\begin{bmatrix}1+2\\-2+5\end{bmatrix}=\begin{bmatrix}3\\3\end{bmatrix}$$

给定向量 $\boldsymbol{u}$ 和实数 c，$\boldsymbol{u}$ 与 c 的**标量乘法**（或**数乘**）是把 $\boldsymbol{u}$ 的每个元素乘以 c，所得向量记为 $c\boldsymbol{u}$. 例如，

$$若\ \boldsymbol{u}=\begin{bmatrix}3\\-1\end{bmatrix},\ c=5,\ 则\ c\boldsymbol{u}=5\begin{bmatrix}3\\-1\end{bmatrix}=\begin{bmatrix}15\\-5\end{bmatrix}$$

$c\boldsymbol{u}$ 中的数 c 称为**标量**（或**数**），它写成细体字母以区别于黑斜体的向量 $\boldsymbol{u}$.

向量加法与标量乘法也可以组合起来，如下例.

例 1 给定 $\boldsymbol{u}=\begin{bmatrix}1\\-2\end{bmatrix}$ 和 $\boldsymbol{v}=\begin{bmatrix}2\\-5\end{bmatrix}$，求 $4\boldsymbol{u}$，$(-3)\boldsymbol{v}$ 以及 $4\boldsymbol{u}+(-3)\boldsymbol{v}$.

解

$$4\boldsymbol{u}=\begin{bmatrix}4\\-8\end{bmatrix},\ (-3)\boldsymbol{v}=\begin{bmatrix}-6\\15\end{bmatrix}$$

而

$$4\boldsymbol{u}+(-3)\boldsymbol{v}=\begin{bmatrix}4\\-8\end{bmatrix}+\begin{bmatrix}-6\\15\end{bmatrix}=\begin{bmatrix}-2\\7\end{bmatrix}$$

■

⊖ 本书大部分内容讨论元素为实数的向量与矩阵. 然而，第 1~5 章的所有定义和定理，以及本书其他部分的大部分内容，对元素是复数时也成立，复向量与矩阵应用于在电气工程和物理中.

有时为了方便（以及节省篇幅），我们把列向量$\begin{bmatrix}3\\-1\end{bmatrix}$写成（3, −1）的形式. 这时，我们用圆括弧表示向量，并在两个元素之间加上逗号，以便区别向量（3, −1）与1×2行矩阵$[3\ \ -1]$，后者使用方括号且两元素之间无逗号.

于是

$$\begin{bmatrix}3\\-1\end{bmatrix}\neq[3\ \ -1]$$

因为这两个矩阵的维度不同，尽管它们有相同的元素.

$\mathbb{R}^2$ 的几何表示

考虑平面上的直角坐标系. 因平面上每个点由实数的有序对确定，我们可把几何点(a,b)与列向量$\begin{bmatrix}a\\b\end{bmatrix}$等同. 所以我们可把$\mathbb{R}^2$看作平面上所有点的集合，见图1-7.

向量$\begin{bmatrix}3\\-1\end{bmatrix}$的几何表示是一条由原点（0, 0）指向点（3, −1）的有向线段（用箭头画出），见图1-8. 在这种情况下，单个的在箭头方向上的点本身并不重要[⊖].

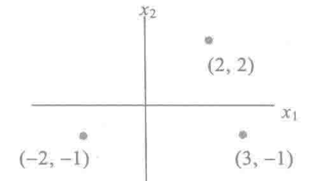


图1-7　用点表示的向量

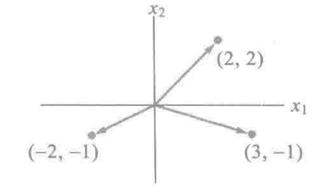
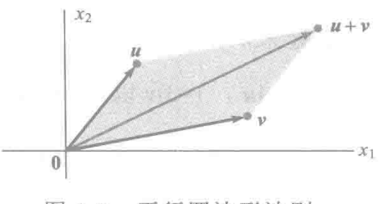


图1-8　用箭头表示向量

向量的和也有很有用的几何意义. 下列规则可用解析几何的知识证明.

> **向量加法的平行四边形法则**
>
> 若$\mathbb{R}^2$中向量$\boldsymbol{u}$和$\boldsymbol{v}$用平面上的点表示，则$\boldsymbol{u}+\boldsymbol{v}$对应于以$\boldsymbol{u}$, $\boldsymbol{0}$和$\boldsymbol{v}$为三个顶点的平行四边形的第4个顶点，见图1-9.
>
> 
>
>
> 图1-9　平行四边形法则

例2　向量$\boldsymbol{u}=\begin{bmatrix}2\\2\end{bmatrix}$, $\boldsymbol{v}=\begin{bmatrix}-6\\1\end{bmatrix}$和$\boldsymbol{u}+\boldsymbol{v}=\begin{bmatrix}-4\\3\end{bmatrix}$，如图1-10所示. ■

⊖ 在物理中，“箭头”可表示力，通常可在空间中自由移动，向量的这种解释将在4.1节讨论.

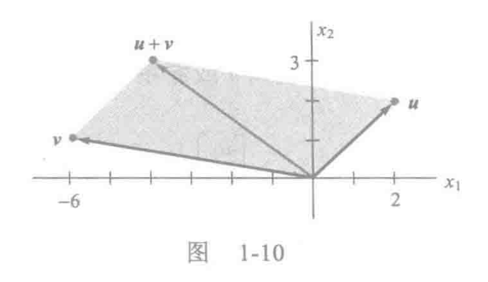


图 1-10

例 3 设 $\boldsymbol{u}=\begin{bmatrix}3\\-1\end{bmatrix}$，在图上表示向量 $\boldsymbol{u}, 2\boldsymbol{u}$和$-\frac{2}{3}\boldsymbol{u}$.

解 见图 1-11，其中表示出向量 $\boldsymbol{u}$，$2\boldsymbol{u}=\begin{bmatrix}6\\-2\end{bmatrix}$和$-\frac{2}{3}\boldsymbol{u}=\begin{bmatrix}-2\\2/3\end{bmatrix}$，$2\boldsymbol{u}$ 表示的箭头长度是 $\boldsymbol{u}$ 表示的箭头长度的 2 倍，指向相同的方向，$-\frac{2}{3}\boldsymbol{u}$ 表示的箭头长度是 $\boldsymbol{u}$ 表示的箭头长度的 2/3 倍且指向相反的方向. 一般地，$c\boldsymbol{u}$ 表示的箭头长度是 $\boldsymbol{u}$ 表示的箭头长度的 $|c|$ 倍. [回忆由（0,0）到 (a,b) 的线段长度为 $\sqrt{a^2+b^2}$，我们将在第 6 章做进一步的讨论.]

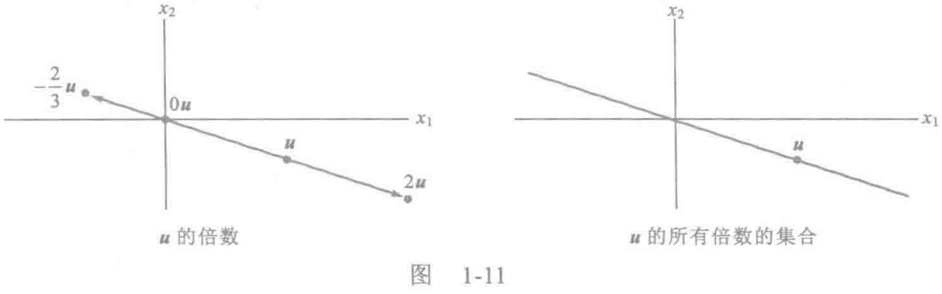


图 1-11 ■

$\mathbb{R}^3$ 中的向量

$\mathbb{R}^3$ 中的向量是 3×1 列矩阵，有 3 个元素. 它们表示 3 维坐标空间中的点，或起点为原点的箭头，图 1-12 表示向量 $\boldsymbol{a}=\begin{bmatrix}2\\3\\4\end{bmatrix}$ 与 $2\boldsymbol{a}$.

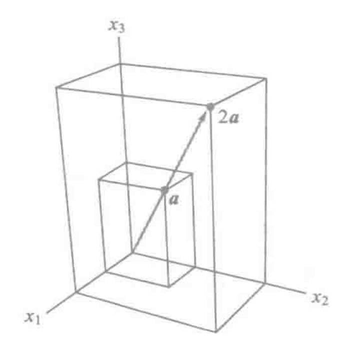


图 1-12 $\mathbb{R}^3$ 中向量的标量乘法

$\mathbb{R}^n$ 中的向量

若 n 是正整数，$\mathbb{R}^n$ 表示所有 n 个实数数列（或有序 n 元组）的集合，通常写成 $n\times1$ 列矩阵的形式，如

$$\boldsymbol{u}=\begin{bmatrix}u_1\\u_2\\\vdots\\u_n\end{bmatrix}$$

所有元素都是零的向量称为零向量，用 $\mathbf{0}$ 表示（$\mathbf{0}$ 中元素的个数可由上下文确定）.

$\mathbb{R}^n$ 中向量相等，以及向量加法与标量乘法运算类似于 $\mathbb{R}^2$ 中的定义. 向量运算有下列性质，它们可直接由实数的相应性质证明. 见本节末的练习题 1 与习题 33 和 34.

> **$\mathbb{R}^n$ 中向量的代数性质**
>
> 对 $\mathbb{R}^n$ 中一切向量 $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}$ 以及标量 c 和 d，
>
> （i）$\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$
>
> （ii）$(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$
>
> （iii）$\boldsymbol{u}+\mathbf{0}=\mathbf{0}+\boldsymbol{u}=\boldsymbol{u}$
>
> （iv）$\boldsymbol{u}+(-\boldsymbol{u})=-\boldsymbol{u}+\boldsymbol{u}=\mathbf{0}$
>
> （v）$c(\boldsymbol{u}+\boldsymbol{v})=c\boldsymbol{u}+c\boldsymbol{v}$
>
> （vi）$(c+d)\boldsymbol{u}=c\boldsymbol{u}+d\boldsymbol{u}$
>
> （vii）$c(d\boldsymbol{u})=(cd)(\boldsymbol{u})$
>
> （viii）$1\boldsymbol{u}=\boldsymbol{u}$
>
> 其中 $-\boldsymbol{u}$ 表示 $(-1)\boldsymbol{u}$.

为了简单起见，我们也使用“向量减法”，用 $\boldsymbol{u}-\boldsymbol{v}$ 代替 $\boldsymbol{u}+(-1)\boldsymbol{v}$，图 1-13 说明 $\boldsymbol{u}-\boldsymbol{v}$ 是 $\boldsymbol{u}$ 和 $-\boldsymbol{v}$ 的和.

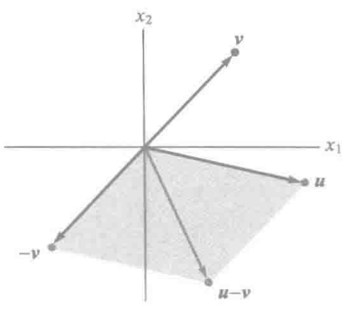


图 1-13 向量减法

线性组合

给定 $\mathbb{R}^n$ 中向量 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p$ 和标量 $c_1,c_2,\cdots,c_p$，向量

$$\boldsymbol{y}=c_1\boldsymbol{v}_1+\cdots+c_p\boldsymbol{v}_p$$

称为向量 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p$ 以 $c_1,c_2,\cdots,c_p$ 为**权的线性组合**. 上述性质(ii)使我们在计算这样的线性组合时不必加上括号. 线性组合中的权可为任意实数，包括零. 例如，下列向量都是 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的线性组合：

$$\sqrt{3}\boldsymbol{v}_1+\boldsymbol{v}_2,\ \frac{1}{2}\boldsymbol{v}_1\left(=\frac{1}{2}\boldsymbol{v}_1+0\boldsymbol{v}_2\right),\ \mathbf{0}(=0\boldsymbol{v}_1+0\boldsymbol{v}_2)$$

例 4 图 1-14 选择性给出向量 $\boldsymbol{v}_1=\begin{bmatrix}-1\\1\end{bmatrix}$ 和 $\boldsymbol{v}_2=\begin{bmatrix}2\\1\end{bmatrix}$ 的某些线性组合.（注：图中平行网格线是通过 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的整数倍画出的.）估计由 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的线性组合生成的向量 $\boldsymbol{u}$ 和 $\boldsymbol{w}$.

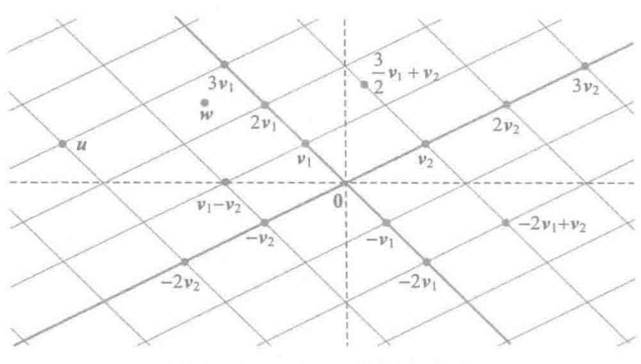


图 1-14　$\boldsymbol{v}_1$ 与 $\boldsymbol{v}_2$ 的线性组合

解　平行四边形法则说明 $\boldsymbol{u}$ 是 $3\boldsymbol{v}_1$ 与 $-2\boldsymbol{v}_2$ 的和，即

$$\boldsymbol{u}=3\boldsymbol{v}_1-2\boldsymbol{v}_2$$

这个 $\boldsymbol{u}$ 的表达式可以解释为经过两条直线路径从原点到达 $\boldsymbol{u}$ 的移动指令. 首先在 $\boldsymbol{v}_1$ 方向移到 3 个单元到达 $3\boldsymbol{v}_1$，然后在 $\boldsymbol{v}_2$ 方向（平行于经过 $\boldsymbol{v}_2$ 和 $\mathbf{0}$ 的直线）移动–2 个单元. 虽然向量 $\boldsymbol{w}$ 不在网格线上，它在两条线的正中间，在由 $(5/2)\boldsymbol{v}_1$ 和 $(-1/2)\boldsymbol{v}_2$ 所确定的平行四边形的顶点处（见图 1-15），因此

$$\boldsymbol{w}=\frac{5}{2}\boldsymbol{v}_1-\frac{1}{2}\boldsymbol{v}_2$$

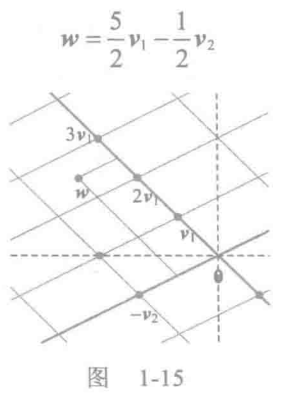


图　1-15

■

下面的例子把线性组合与 1.1 节、1.2 节的存在性问题联系起来.

例 5　设 $\boldsymbol{a}_1=\begin{bmatrix}1\\-2\\-5\end{bmatrix}$，$\boldsymbol{a}_2=\begin{bmatrix}2\\5\\6\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}7\\4\\-3\end{bmatrix}$，确定 $\boldsymbol{b}$ 能否写成 $\boldsymbol{a}_1$ 和 $\boldsymbol{a}_2$ 的线性组合，也就是说，确定是否存在权 x_1 和 x_2 使

$$x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2=\boldsymbol{b} \tag{1}$$

若向量方程（1）有解，求它的解.

解 根据向量加法和标量乘法的定义把向量方程

$$x_1\begin{bmatrix}1\\-2\\-5\end{bmatrix}+x_2\begin{bmatrix}2\\5\\6\end{bmatrix}=\begin{bmatrix}7\\4\\-3\end{bmatrix}$$

$$\uparrow\ \boldsymbol{a}_1 \qquad \uparrow\ \boldsymbol{a}_2 \qquad \uparrow\ \boldsymbol{b}$$

写成

$$\begin{bmatrix}x_1\\-2x_1\\-5x_1\end{bmatrix}+\begin{bmatrix}2x_2\\5x_2\\6x_2\end{bmatrix}=\begin{bmatrix}7\\4\\-3\end{bmatrix}$$

或

$$\begin{bmatrix}x_1+2x_2\\-2x_1+5x_2\\-5x_1+6x_2\end{bmatrix}=\begin{bmatrix}7\\4\\-3\end{bmatrix} \tag{2}$$

（2）式左右两边的向量相等当且仅当它们的对应元素相等. 即 x_1 和 x_2 满足向量方程（1），当且仅当 x_1 和 x_2 满足方程组

$$\begin{aligned}x_1+2x_2&=7\\-2x_1+5x_2&=4\\-5x_1+6x_2&=-3\end{aligned} \tag{3}$$

我们用行化简算法将上述线性方程组的增广矩阵化简，以此解方程组[㊀]：

$$\begin{bmatrix}1&2&7\\-2&5&4\\-5&6&-3\end{bmatrix}\sim\begin{bmatrix}1&2&7\\0&9&18\\0&16&32\end{bmatrix}\sim\begin{bmatrix}1&2&7\\0&1&2\\0&16&32\end{bmatrix}\sim\begin{bmatrix}1&0&3\\0&1&2\\0&0&0\end{bmatrix}$$

（3）的解是 $x_1=3, x_2=2$，因此 $\boldsymbol{b}$ 是 $\boldsymbol{a}_1$ 与 $\boldsymbol{a}_2$ 的线性组合，权为 $x_1=3$ 和 $x_2=2$，即

$$3\begin{bmatrix}1\\-2\\-5\end{bmatrix}+2\begin{bmatrix}2\\5\\6\end{bmatrix}=\begin{bmatrix}7\\4\\-3\end{bmatrix}$$ ■

注意例 5 中原来的向量 $\boldsymbol{a}_1$， $\boldsymbol{a}_2$ 和 $\boldsymbol{b}$ 是我们进行行化简的增广矩阵的列

$$\begin{bmatrix}1&2&7\\-2&5&4\\-5&6&-3\end{bmatrix}$$

$$\uparrow\ \boldsymbol{a}_1 \quad \uparrow\ \boldsymbol{a}_2 \quad \uparrow\ \boldsymbol{b}$$

让我们将此矩阵写成另一形式，使它的列备受关注，即

$$[\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \quad \boldsymbol{b}] \tag{4}$$

这样，由向量方程（1）可以直接写出增广矩阵而不必经过例 5 中的中间步骤. 就是按它们在（1）中出现的次序排列，就得到矩阵（4）.

㊀ 记号“~”表示矩阵行等价（见 1.2 节）.

由上述讨论我们可以得到以下结论.

> 向量方程
> $$x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+\cdots+x_n\boldsymbol{a}_n=\boldsymbol{b}$$
> 和增广矩阵为
> $$[\boldsymbol{a}_1\quad \boldsymbol{a}_2\quad \cdots\quad \boldsymbol{a}_n\quad \boldsymbol{b}] \tag{5}$$
> 的线性方程组有相同的解集. 特别地, $\boldsymbol{b}$ 可表示为 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$ 的线性组合, 当且仅当对应于(5)式的方程组有解.

线性代数的一个主要思想是研究可以表示为某一固定向量集合 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p\}$ 的线性组合的所有向量.

定义 若 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p$ 是 $\mathbb{R}^n$ 中的向量, 则 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p$ 的所有线性组合所成的集合用记号 $\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p\}$ 表示, 称为由 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p$ 所**生成(或张成)的 $\mathbb{R}^n$ 的子集**, 也就是说, $\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p\}$ 是所有形如
$$c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_p\boldsymbol{v}_p$$
的向量的集合, 其中 $c_1,c_2,\cdots,c_p$ 为标量.

要判断向量 $\boldsymbol{b}$ 是否属于 $\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p\}$, 就是判断向量方程
$$x_1\boldsymbol{v}_1+x_2\boldsymbol{v}_2+\cdots+x_p\boldsymbol{v}_p=\boldsymbol{b}$$
是否有解, 或等价地, 判断增广矩阵为 $[\boldsymbol{v}_1\ \boldsymbol{v}_2\cdots\boldsymbol{v}_p\ \boldsymbol{b}]$ 的线性方程组是否有解.

注意 $\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p\}$ 包含 $\boldsymbol{v}_1$ 的所有倍数, 因 $c\boldsymbol{v}_1=c\boldsymbol{v}_1+0\boldsymbol{v}_2+\cdots+0\boldsymbol{v}_p$, 特别地, 它一定包含零向量.

$\text{Span}\{\boldsymbol{v}\}$ 与 $\text{Span}\{\boldsymbol{u},\boldsymbol{v}\}$ 的几何解释

设 $\boldsymbol{v}$ 是 $\mathbb{R}^3$ 中的向量, 那么 $\text{Span}\{\boldsymbol{v}\}$ 就是 $\boldsymbol{v}$ 的所有数量倍数的集合, 也就是通过 $\boldsymbol{v}$ 和 $\boldsymbol{0}$ 的直线上的所有点的集合, 见图 1-16.

若 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是 $\mathbb{R}^3$ 中的非零向量, $\boldsymbol{v}$ 不是 $\boldsymbol{u}$ 的倍数, 则 $\text{Span}\{\boldsymbol{u},\boldsymbol{v}\}$ 是 $\mathbb{R}^3$ 中通过 $\boldsymbol{u}$, $\boldsymbol{v}$ 和 $\boldsymbol{0}$ 的平面, 特别地, $\text{Span}\{\boldsymbol{u},\boldsymbol{v}\}$ 包含 $\mathbb{R}^3$ 中通过 $\boldsymbol{u}$ 与 $\boldsymbol{0}$ 的直线, 也包含通过 $\boldsymbol{v}$ 与 $\boldsymbol{0}$ 的直线(见图 1-17).

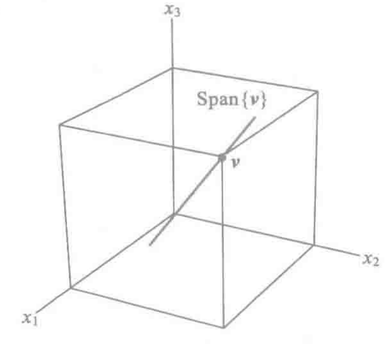


图 1-16 $\text{Span}\{\boldsymbol{v}\}$ 是通过原点的直线

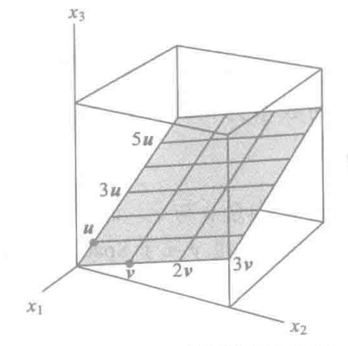


图 1-17 $\text{Span}\{\boldsymbol{u},\boldsymbol{v}\}$ 是通过原点的平面

例 6 设 $\boldsymbol{a}_1=\begin{bmatrix}1\\-2\\3\end{bmatrix}$, $\boldsymbol{a}_2=\begin{bmatrix}5\\-13\\-3\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}-3\\8\\1\end{bmatrix}$，则 $\mathrm{Span}\{\boldsymbol{a}_1,\boldsymbol{a}_2\}$ 是 $\mathbb{R}^3$ 中通过原点的一个平面，问 $\boldsymbol{b}$ 是否在该平面内？

解 方程 $x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2=\boldsymbol{b}$ 是否有解？为回答此问题，把增广矩阵 $[\boldsymbol{a}_1\quad \boldsymbol{a}_2\quad \boldsymbol{b}]$ 进行行化简

$$\begin{bmatrix}1&5&-3\\-2&-13&8\\3&-3&1\end{bmatrix}\sim\begin{bmatrix}1&5&-3\\0&-3&2\\0&-18&10\end{bmatrix}\sim\begin{bmatrix}1&5&-3\\0&-3&2\\0&0&-2\end{bmatrix}$$

第 3 个方程为 $0x_2=-2$，它说明方程组无解，向量方程 $x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2=\boldsymbol{b}$ 无解，故 $\boldsymbol{b}$ 不属于 $\mathrm{Span}\{\boldsymbol{a}_1,\boldsymbol{a}_2\}$．■

应用中的线性组合

本节最后一个例子说明标量乘法与线性组合在某一个量，例如“成本”，被分解成若干部分时的一个应用. 这个例子的基本原理是，生产某种产品的单位成本已知时，可求出生产若干个单位的成本.

$$\{单位数量\}\times\{每单位成本\}=\{总成本\}$$

例 7 某公司生产两种产品，每美元价值的产品 B，公司需耗费 0.45 美元材料，0.25 美元劳动，0.15 美元管理费用，对每美元价值的产品 C，公司耗费 0.40 美元材料，0.30 美元劳动，0.15 美元管理费用，设

$$\boldsymbol{b}=\begin{bmatrix}0.45\\0.25\\0.15\end{bmatrix}\quad \boldsymbol{c}=\begin{bmatrix}0.40\\0.30\\0.15\end{bmatrix}$$

则 $\boldsymbol{b}$ 和 $\boldsymbol{c}$ 称为两种产品的“单位美元产出成本”.

a．向量 $100\boldsymbol{b}$ 的经济解释是什么?

b．设公司希望生产 x_1 美元产品 B 和 x_2 美元产品 C．给出描述该公司花费的各部分成本（材料，劳动，管理费用）的向量.

解 a．我们有

$$100\boldsymbol{b}=100\begin{bmatrix}0.45\\0.25\\0.15\end{bmatrix}=\begin{bmatrix}45\\25\\15\end{bmatrix}$$

向量 $100\boldsymbol{b}$ 列出生产 100 美元的产品 B 需要的各种成本，即 45 美元材料、25 美元劳动、15 美元管理费用.

b. 生产 x_1 美元的产品 B 的成本由向量 $x_1\boldsymbol{b}$ 给出，生产 x_2 美元的产品 C 的成本由向量 $x_2\boldsymbol{c}$ 给出. 因此总的成本为 $x_1\boldsymbol{b}+x_2\boldsymbol{c}$． ■

练习题

1. 对 $\mathbb{R}^n$ 中的任意向量 $\boldsymbol{u}$ 与 $\boldsymbol{v}$，证明 $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$.
2. h 取什么值时，向量 $\boldsymbol{y}$ 属于 $\mathrm{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$，若

$$\boldsymbol{v}_1=\begin{bmatrix}1\\-1\\-2\end{bmatrix},\ \boldsymbol{v}_2=\begin{bmatrix}5\\-4\\-7\end{bmatrix},\ \boldsymbol{v}_3=\begin{bmatrix}-3\\1\\0\end{bmatrix},\ \boldsymbol{y}=\begin{bmatrix}-4\\3\\h\end{bmatrix}$$

习题 1.3

在习题 1 和 2 中，计算 $\boldsymbol{u}+\boldsymbol{v}$ 与 $\boldsymbol{u}-2\boldsymbol{v}$.

1. $\boldsymbol{u}=\begin{bmatrix}-1\\2\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}-3\\-1\end{bmatrix}$

2. $\boldsymbol{u}=\begin{bmatrix}3\\2\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}2\\-1\end{bmatrix}$

在习题 3 和 4 中，用箭头在 xy 平面上表示下列向量：$\boldsymbol{u},\boldsymbol{v},-\boldsymbol{v},-2\boldsymbol{v},\boldsymbol{u}+\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v},\boldsymbol{u}-2\boldsymbol{v}$，注意 $\boldsymbol{u}-\boldsymbol{v}$ 是三个顶点为 $\boldsymbol{u},\boldsymbol{0}$ 和 $-\boldsymbol{v}$ 的平行四边形的另一个顶点.

3. $\boldsymbol{u}$ 与 $\boldsymbol{v}$ 如习题 1.

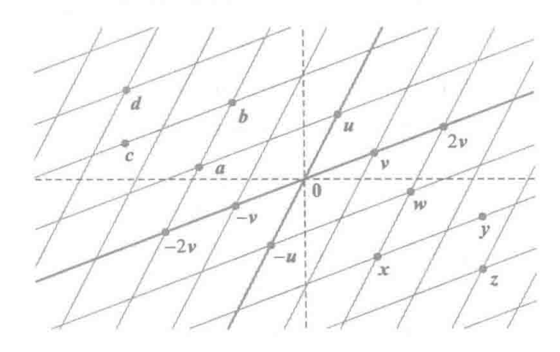
4. $\boldsymbol{u}$ 与 $\boldsymbol{v}$ 如习题 2.

在习题 5 和 6 中，写出等价于所给向量方程的线性方程组.

5. $x_1\begin{bmatrix}3\\-2\\8\end{bmatrix}+x_2\begin{bmatrix}5\\0\\-9\end{bmatrix}=\begin{bmatrix}2\\-3\\8\end{bmatrix}$

6. $x_1\begin{bmatrix}3\\-2\end{bmatrix}+x_2\begin{bmatrix}7\\3\end{bmatrix}+x_3\begin{bmatrix}-2\\1\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$

利用附带的图把习题 7 和 8 中的向量用 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的线性组合表示. $\mathbb{R}^2$ 中的任意向量是否一定是向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的线性组合？



7. 向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d}$

8. 向量 $\boldsymbol{w},\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}$

在习题 9 和 10 中写出等价于给出的方程组的向量方程：

9.
$$\begin{aligned} x_2+5x_3&=0\\ 4x_1+6x_2-x_3&=0\\ -x_1+3x_2-8x_3&=0\end{aligned}$$

10.
$$\begin{aligned} 3x_1-2x_2+4x_3&=3\\ -2x_1-7x_2+5x_3&=1\\ 5x_1+4x_2-3x_3&=2\end{aligned}$$

在习题 11 和 12 中，确定 $\boldsymbol{b}$ 是否是 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 的线性组合.

11. $\boldsymbol{a}_1=\begin{bmatrix}1\\-2\\0\end{bmatrix}$，$\boldsymbol{a}_2=\begin{bmatrix}0\\1\\2\end{bmatrix}$，$\boldsymbol{a}_3=\begin{bmatrix}5\\-6\\8\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}2\\-1\\6\end{bmatrix}$

12. $\boldsymbol{a}_1=\begin{bmatrix}1\\0\\1\end{bmatrix}$，$\boldsymbol{a}_2=\begin{bmatrix}-2\\3\\-2\end{bmatrix}$，$\boldsymbol{a}_3=\begin{bmatrix}-6\\7\\5\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}11\\-5\\9\end{bmatrix}$

在习题 13 和 14 中，确定 $\boldsymbol{b}$ 是否矩阵 A 的各列向量的线性组合.

13. $A=\begin{bmatrix}1&-4&2\\0&3&5\\-2&8&-4\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}3\\-7\\-3\end{bmatrix}$

14. $A=\begin{bmatrix}1&0&5\\-2&1&-6\\0&2&8\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}2\\-1\\6\end{bmatrix}$

15. 设 $\boldsymbol{a}_1=\begin{bmatrix}1\\3\\-1\end{bmatrix}$，$\boldsymbol{a}_2=\begin{bmatrix}-5\\-8\\2\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}3\\-5\\h\end{bmatrix}$，当 h 取何值时 $\boldsymbol{b}$ 是在 $\boldsymbol{a}_1$ 和 $\boldsymbol{a}_2$ 生成的平面内？

16. 设 $\boldsymbol{v}_1=\begin{bmatrix}1\\0\\-2\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}-2\\1\\7\end{bmatrix}$，$\boldsymbol{y}=\begin{bmatrix}h\\-3\\-5\end{bmatrix}$，当 h 取何值时 $\boldsymbol{y}$ 在 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 生成的平面内？

在习题 17 和 18 中，写出属于 $\mathrm{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 中的 5 个向量以及用来生成这些向量的 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 的权，并写出这些向量的 3 个元素. 不用画图.

17. $\boldsymbol{v}_1=\begin{bmatrix}3\\1\\2\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}-4\\0\\1\end{bmatrix}$

18. $\boldsymbol{v}_1=\begin{bmatrix}1\\1\\-2\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}-2\\3\\0\end{bmatrix}$

19. 对下面的向量，给出 $\mathrm{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 的几何解释.

$$\boldsymbol{v}_1=\begin{bmatrix}8\\2\\-6\end{bmatrix},\quad \boldsymbol{v}_2=\begin{bmatrix}12\\3\\-9\end{bmatrix}$$

20. 对习题 18 中的向量，给出 $\mathrm{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 的几何解释.

21. 设 $u=\begin{bmatrix}2\\-1\end{bmatrix}$，$v=\begin{bmatrix}2\\1\end{bmatrix}$，证明：对所有 h 和 k，$\begin{bmatrix}h\\k\end{bmatrix}$ 属于 Span$\{u,v\}$.

22. 构造一 3×3 无零元素的矩阵 A 和一 $\mathbb{R}^3$ 中的向量 b，使得 b 不属于 A 的列向量的生成集.

在习题 23 和 24 中，判断下列命题的真假，给出理由.

23. a. 向量 $\begin{bmatrix}-4\\3\end{bmatrix}$ 的另一种写法是 $[-4\quad 3]$.

b. 平面上对应于 $\begin{bmatrix}-2\\5\end{bmatrix}$ 和 $\begin{bmatrix}-5\\2\end{bmatrix}$ 的两点位于通过原点的一条直线上.

c. 向量 $\frac{1}{2}v_1$ 是向量 v_1 和 v_2 的线性组合.

d. 增广矩阵为 $[a_1\quad a_2\quad a_3\quad b]$ 的线性方程组的解集与向量方程 $x_1a_1+x_2a_2+x_3a_3=b$ 的解集相同.

e. 集 Span$\{u,v\}$ 总是表示通过原点的一个平面.

24. a. 当 u 和 v 是非零向量时，Span$\{u,v\}$ 包含通过 u 与原点的直线.

b. 任意 5 个实数组成的数列是 $\mathbb{R}^5$ 中一个向量.

c. 对应增广矩阵 $[a_1\quad a_2\quad a_3\quad b]$ 的线性方程组是否有解的问题等价于 b 是否属于 Span$\{a_1, a_2, a_3\}$ 的问题.

d. 向量 u 等于向量 $u-v$ 与向量 v 之和.

e. 在线性组合 $c_1v_1+c_2v_2+\cdots+c_pv_p$ 中，权 $c_1, c_2,\cdots,c_p$ 不能全为 0.

25. 设 $A=\begin{bmatrix}1&0&-4\\0&3&-2\\-2&6&3\end{bmatrix}$，$b=\begin{bmatrix}4\\1\\-4\end{bmatrix}$，以 a_1,a_2,a_3 表示 A 的各列. 并设 $W=\text{Span}\{a_1,a_2,a_3\}$.

a. b 是否属于 $\{a_1, a_2, a_3\}$？在 $\{a_1, a_2, a_3\}$ 中有多少个向量？

b. b 是否属于 W？W 中有多少个向量？

c. 证明：a_1 属于 W（提示：不必作行变换）.

26. 设 $A=\begin{bmatrix}2&0&6\\-1&8&5\\1&-2&1\end{bmatrix}$，$b=\begin{bmatrix}10\\3\\7\end{bmatrix}$，设 W 为 A 的列向量的所有线性组合的集合.

a. b 是否属于 W？

b. 证明：A 的第 2 列属于 W.

27. 某矿业公司有两个矿，#1 矿每天生产 30 吨铜和 600 公斤银，#2 矿每天生产 40 吨铜和 380 公斤银，设 $v_1=\begin{bmatrix}30\\600\end{bmatrix}$，$v_2=\begin{bmatrix}40\\380\end{bmatrix}$，则 v_1 和 v_2 分别表示#1 矿与#2 矿的“日产出”.

a. 向量 $5v_1$ 有什么实际意义？

b. 设该公司的#1 矿生产 x_1 天，#2 矿生产 x_2 天，写出表示该公司生产 240 吨铜与 2 824 公斤银时，各矿生产天数的向量方程，不必解方程.

c. [M]解（b）中的方程.

28. 某蒸汽厂烧两种煤：无烟煤（A）和烟煤（B），每吨煤 A 燃烧产生 27.6 百万焦耳的热量，3 100 克二氧化硫和 250 克固体粒子污染物. 每吨煤 B 燃烧产生 30.2 百万焦耳的热量，6 400 克二氧化硫和 360 克固体粒子污染物.

a. 若该厂燃烧 x_1 吨煤 A 和 x_2 吨煤 B，它产出多少热量？

b. 设该厂的产出可用它产出的热量，二氧化硫和固体粒子污染物的量构成的向量表示. 把这个产出用两个向量的线性组合表示，设它燃烧煤 A x_1 吨和煤 B x_2 吨.

c. [M]设某一段时间，该厂生产 162 百万焦耳的热量，23 610 克二氧化硫和 1 623 克固体粒子污染物，写出向量方程，并确定该厂烧了两种煤各多少吨？

29. 设 $v_1, v_2,\cdots,v_k$ 是 $\mathbb{R}^3$ 的点，且对 $j=1,2,\cdots,k$ 在点 v_j 有质量为 m_j 的物体. 物理学家称这些物体为质点. 这个质点系的总质量为

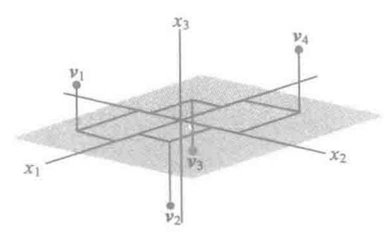
$$m=m_1+m_2+\cdots+m_k$$

质点系的重心（或质心）是

$$\overline{v}=\frac{1}{m}(m_1v_1+m_2v_2+\cdots+m_kv_k)$$

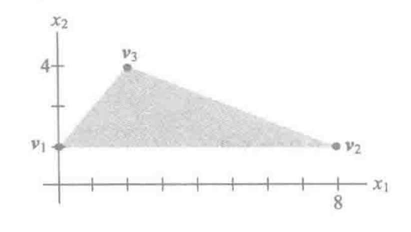
计算由下列质点组成的质点系的重心（见下图）：

点	质量
$v_1=(2,-2,4)$	4 克
$v_2=(-4,2,3)$	2 克
$v_3=(4,0,-2)$	3 克
$v_4=(1,-6,0)$	5 克



30. 设 $\boldsymbol{v}$ 为习题 29 中的一组质点 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k$ 的质心，$\boldsymbol{v}$ 是否属于 $\text{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k\}$？给出理由.

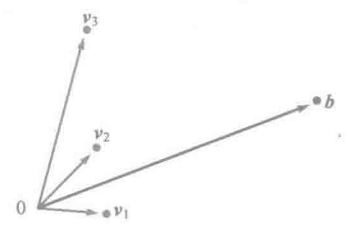
31. 一块密度和厚度均匀的三角形薄板，其三个顶点分别是 $\boldsymbol{v}_1=(0,1),\boldsymbol{v}_2=(8,1),\boldsymbol{v}_3=(2,4)$，如下图所示，且其质量为 3 克.



a. 求这块薄板的质心坐标 (x,y).这块薄板的“平衡点”与在三个顶点上有 1 克质量的薄板的质心相重合.

b. 求如何在三个顶点上分布额外的 6 克物质，使得薄板的平衡点定位在点（2,2）上.（提示：设 w_1,w_2,w_3 分别表示加在三个顶点上的质量，则 $w_1+w_2+w_3=6$.）

32. 考虑在 $\mathbb{R}^2$ 中的向量 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$和$\boldsymbol{b}$，如下图所示，方程 $x_1\boldsymbol{v}_1+x_2\boldsymbol{v}_2+x_3\boldsymbol{v}_3=\boldsymbol{b}$ 是否有解？解是否唯一？使用图形给出解释.



33. 设 $\boldsymbol{u}=(u_1,u_2,\cdots,u_n)$，$\boldsymbol{v}=(v_1,v_2,\cdots,v_n)$，$\boldsymbol{w}=(w_1,w_2,\cdots,w_n)$，证明 $\mathbb{R}^n$ 的下列代数性质：
   a. $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$.
   b. $c(\boldsymbol{u}+\boldsymbol{v})=c\boldsymbol{u}+c\boldsymbol{v}$，其中$c$ 为任意数.

34. 设 $\boldsymbol{u}=(u_1,u_2,\cdots,u_n)$，证明 $\mathbb{R}^n$ 的下列代数性质：
   a. $\boldsymbol{u}+(-\boldsymbol{u})=(-\boldsymbol{u})+\boldsymbol{u}=\boldsymbol{0}$.
   b. $c(d\boldsymbol{u})=(cd)\boldsymbol{u}$，其中$c,d$ 为任意数.

练习题答案

1. 取 $\mathbb{R}^n$ 中任意向量 $\boldsymbol{u}=(u_1,\ u_2,\cdots,u_n)$，$\boldsymbol{v}=(v_1,v_2,\cdots,v_n)$，

$$\begin{aligned}\boldsymbol{u}+\boldsymbol{v}&=(u_1+v_1,\ u_2+v_2,\cdots,\ u_n+v_n) && \text{向量加法的定义}\\&=(v_1+u_1,\ v_2+u_2,\cdots,v_n+u_n) && \mathbb{R}\text{中加法交换律}\\&=\boldsymbol{v}+\boldsymbol{u} && \text{向量加法的定义}\end{aligned}$$

2. 向量 $\boldsymbol{y}$ 属于 $\text{Span}\{\boldsymbol{v}_1,\ \boldsymbol{v}_2,\ \boldsymbol{v}_3\}$，当且仅当存在数 $x_1,\ x_2,\ x_3$ 使

$$x_1\begin{bmatrix}1\\-1\\-2\end{bmatrix}+x_2\begin{bmatrix}5\\-4\\-7\end{bmatrix}+x_3\begin{bmatrix}-3\\1\\0\end{bmatrix}=\begin{bmatrix}-4\\3\\h\end{bmatrix}$$

该向量方程等价于含 3 个未知数的 3 个方程组成的方程组. 化简这个方程组的增广矩阵，得

$$\begin{bmatrix} 1 & 5 & -3 & -4 \\ -1 & -4 & 1 & 3 \\ -2 & -7 & 0 & h \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 & -4 \\ 0 & 1 & -2 & -1 \\ 0 & 3 & -6 & h-8 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 & -4 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & h-5 \end{bmatrix}$$

当且仅当第 4 列没有主元时，方程组是相容的，也就是说，$h-5$ 必须为 0，因此当且仅当 $h=5$ 时，$\boldsymbol{y}$ 属于 Span$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$. 见图 1-18.

记住：方程组中有自由变量并不能保证方程组有解.

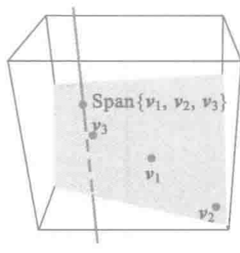


图 1-18 当 $h=5$ 时，点 $\begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$ 位于与平面相交的直线上

1.4 矩阵方程 $A\boldsymbol{x}=\boldsymbol{b}$

线性代数中的一个基本思想是把向量的线性组合看作矩阵与向量的积. 下列定义允许我们将 1.3 节的某些概念用新的方法表述出来.

定义 若 A 是 $m\times n$ 矩阵，它的各列为 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$. 若 $\boldsymbol{x}$ 是 $\mathbb{R}^n$ 中向量，则 A 与 $\boldsymbol{x}$ **的积**，记为 $A\boldsymbol{x}$，就是 A **的各列以 $\boldsymbol{x}$ 中对应元素为权的线性组合**，即

$$A\boldsymbol{x}=[\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \quad \cdots \quad \boldsymbol{a}_n]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}=x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+\cdots+x_n\boldsymbol{a}_n$$

注意 $A\boldsymbol{x}$ 仅当 A 的列数等于 $\boldsymbol{x}$ 中元素个数时才有定义.

例 1

a. $\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix}=4\begin{bmatrix} 1 \\ 0 \end{bmatrix}+3\begin{bmatrix} 2 \\ -5 \end{bmatrix}+7\begin{bmatrix} -1 \\ 3 \end{bmatrix}=\begin{bmatrix} 4 \\ 0 \end{bmatrix}+\begin{bmatrix} 6 \\ -15 \end{bmatrix}+\begin{bmatrix} -7 \\ 21 \end{bmatrix}=\begin{bmatrix} 3 \\ 6 \end{bmatrix}$

b. $\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix}\begin{bmatrix} 4 \\ 7 \end{bmatrix}=4\begin{bmatrix} 2 \\ 8 \\ -5 \end{bmatrix}+7\begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix}=\begin{bmatrix} 8 \\ 32 \\ -20 \end{bmatrix}+\begin{bmatrix} -21 \\ 0 \\ 14 \end{bmatrix}=\begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}$ ■

例 2 对 $\mathbb{R}^m$ 中 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$，把线性组合 $3\boldsymbol{v}_1-5\boldsymbol{v}_2+7\boldsymbol{v}_3$ 表示为矩阵乘向量的形式.

解 把 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 排列成矩阵 A，把数 3，−5，7 排列成向量 $\boldsymbol{x}$，即

$$3\boldsymbol{v}_1-5\boldsymbol{v}_2+7\boldsymbol{v}_3=[\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \boldsymbol{v}_3]\begin{bmatrix} 3 \\ -5 \\ 7 \end{bmatrix}=A\boldsymbol{x}$$

■

在 1.3 节中我们学习了将线性方程组写成包含向量线性组合形式的向量方程，例如，方程组

$$\begin{aligned} x_1+2x_2-\ x_3&=4 \\ -5x_2+3x_3&=1 \end{aligned} \tag{1}$$

等价于

$$x_1\begin{bmatrix}1\\0\end{bmatrix}+x_2\begin{bmatrix}2\\-5\end{bmatrix}+x_3\begin{bmatrix}-1\\3\end{bmatrix}=\begin{bmatrix}4\\1\end{bmatrix} \tag{2}$$

如例 2 所示，我们也可将方程左边的线性组合写成矩阵乘向量的形式，（2）成为

$$\begin{bmatrix}1&2&-1\\0&-5&3\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}4\\1\end{bmatrix} \tag{3}$$

方程（3）有形式 $A\boldsymbol{x}=\boldsymbol{b}$，我们称这样的方程为**矩阵方程**，以区别于（2）式那样的向量方程.

注意（3）中的矩阵仅是方程（1）中的系数矩阵，类似的计算说明，任何线性方程组，或类似于（2）式的向量方程，可以写成等价的形式为 $A\boldsymbol{x}=\boldsymbol{b}$ 的矩阵方程，这一简单的观点将在本书中重复使用.

下面是正式的结果.

定理 3 若 A 是 $m\times n$ 矩阵，它的各列为 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$，而 $\boldsymbol{b}$ 属于 $\mathbb{R}^m$，则矩阵方程

$$A\boldsymbol{x}=\boldsymbol{b} \tag{4}$$

与向量方程

$$x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+\cdots+x_n\boldsymbol{a}_n=\boldsymbol{b} \tag{5}$$

有相同的解集. 它又与增广矩阵为

$$\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n & \boldsymbol{b}\end{bmatrix} \tag{6}$$

的线性方程组有相同的解集.

定理 3 给出了研究线性代数问题的一个有力工具，使我们现在可将线性方程组用三种不同但彼此等价的观点来研究：作为矩阵方程、作为向量方程或作为线性方程组，当我们构造实际生活中某个问题的数学模型时，我们可自由地选择任何一种最自然的观点. 于是我们可在方便的时候由一种观点转向另一种观点. 任何情况下，矩阵方程、向量方程以及线性方程组都用相同方法来解——即用行化简算法来化简增广矩阵（6）. 其他解法将在以后讨论.

解的存在性

$A\boldsymbol{x}$ 的定义直接导致下列有用的事实.

> 方程 $A\boldsymbol{x}=\boldsymbol{b}$ 有解当且仅当 $\boldsymbol{b}$ 是 A 的各列的线性组合.

在 1.3 节中，我们考虑了存在性问题，“$\boldsymbol{b}$ 是否属于 $\mathrm{Span}\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n\}$？”等价地，“$A\boldsymbol{x}=\boldsymbol{b}$ 是否相容？”一个更困难的问题是要确定方程 $A\boldsymbol{x}=\boldsymbol{b}$ 对任意的 $\boldsymbol{b}$ 是否有解.

例 3 设 $A=\begin{bmatrix}1&3&4\\-4&2&-6\\-3&-2&-7\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}$. 方程 $A\boldsymbol{x}=\boldsymbol{b}$ 是否对一切可能的 b_1,b_2,b_3 有解？

解 把 $\boldsymbol{Ax}=\boldsymbol{b}$ 的增广矩阵行化简，

$$\begin{bmatrix} 1 & 3 & 4 & b_1 \\ -4 & 2 & -6 & b_2 \\ -3 & -2 & -7 & b_3 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & b_2+4b_1 \\ 0 & 7 & 5 & b_3+3b_1 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & b_2+4b_1 \\ 0 & 0 & 0 & b_3+3b_1-\frac{1}{2}(b_2+4b_1) \end{bmatrix}$$

增广列的第 3 个元素为 $b_1-\frac{1}{2}b_2+b_3$，方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 并不是对一切的 $\boldsymbol{b}$ 都有解，因 $b_1-\frac{1}{2}b_2+b_3$ 可能不为零. ■

例 3 中的简化矩阵描述了使 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解的所有 $\boldsymbol{b}$ 的集合；$\boldsymbol{b}$ 必须满足

$$b_1-\frac{1}{2}b_2+b_3=0$$

这是 $\mathbb{R}^3$ 中一个通过原点的平面，这个平面就是 A 的 3 列的所有线性组合的集合，见图 1-19.

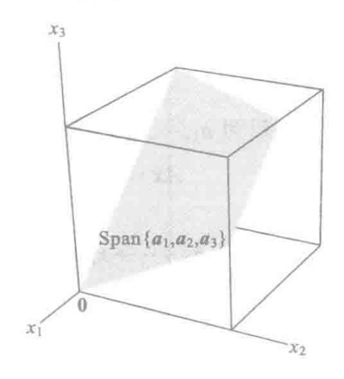


图 1-19 $A=[\boldsymbol{a}_1\ \ \boldsymbol{a}_2\ \ \boldsymbol{a}_3]$ 的各列生成通过原点的平面

例 3 中的方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 并非对所有的 $\boldsymbol{b}$ 都相容，这是因为 A 的阶梯形含有零行. 假如 A 在所有三行都有主元，我们就不必注意增广列的计算，因为这时增广矩阵的阶梯形状不可能产生如 $[0\ \ 0\ \ 0\ \ 1]$ 的行.

在下一定理中，当我们说 **A 的列生成** $\mathbb{R}^m$ 时，意思是说 $\mathbb{R}^m$ 中的每个向量 $\boldsymbol{b}$ 都是 A 的列的线性组合. 一般地，$\mathbb{R}^m$ 中向量集 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ **生成** $\mathbb{R}^m$，意思是说，$\mathbb{R}^m$ 中的每个向量都是 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 的线性组合，即 $\text{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}=\mathbb{R}^m$.

定理 4 设 A 是 $m\times n$ 矩阵，则下列命题是逻辑上等价的，也就是说，对某个 A，它们都成立或者都不成立.

a. 对 $\mathbb{R}^m$ 中每个 $\boldsymbol{b}$，方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解.

b. $\mathbb{R}^m$ 中的每个 $\boldsymbol{b}$ 都是 A 的列的一个线性组合.

c. A 的各列生成 $\mathbb{R}^m$.

d. A 在每一行都有一个主元位置.

定理 4 是本章中最有用的定理之一. 命题（a)、(b）和（c）等价是根据 $A\mathbf{x}$ 的定义和一组向量生成 $\mathbb{R}^m$ 空间的含意而得到的. 再根据例 3 后面的讨论得到命题（a）和（d）等价. 定理完整的证明在本节末尾中给出. 在习题中给出应用定理 4 的例子.

警告 定理 4 讨论的是系数矩阵，而非增广矩阵，若增广矩阵 $[A\ \ \mathbf{b}]$ 在每一行都有主元位置，方程 $A\mathbf{x}=\mathbf{b}$ 可能相容，也可能不相容.

$A\mathbf{x}$ 的计算

例 1 中的计算是根据矩阵 A 与向量 $\mathbf{x}$ 的乘积的定义. 下面的例子给出手工计算 $A\mathbf{x}$ 的元素的一种更有效的方法.

例 4 计算 $A\mathbf{x}$，其中 $A=\begin{bmatrix}2 & 3 & 4\\ -1 & 5 & -3\\ 6 & -2 & 8\end{bmatrix}$，$\mathbf{x}=\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$.

解 由定义

$$\begin{bmatrix}2 & 3 & 4\\ -1 & 5 & -3\\ 6 & -2 & 8\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=x_1\begin{bmatrix}2\\ -1\\ 6\end{bmatrix}+x_2\begin{bmatrix}3\\ 5\\ -2\end{bmatrix}+x_3\begin{bmatrix}4\\ -3\\ 8\end{bmatrix}$$
$$=\begin{bmatrix}2x_1\\ -x_1\\ 6x_1\end{bmatrix}+\begin{bmatrix}3x_2\\ 5x_2\\ -2x_2\end{bmatrix}+\begin{bmatrix}4x_3\\ -3x_3\\ 8x_3\end{bmatrix}=\begin{bmatrix}2x_1+3x_2+4x_3\\ -x_1+5x_2-3x_3\\ 6x_1-2x_2+8x_3\end{bmatrix} \tag{7}$$

矩阵 $A\mathbf{x}$ 的第一个元素是 A 的第一行与 $\mathbf{x}$ 中相应元素乘积之和（有时称为点积），即

$$\begin{bmatrix}2 & 3 & 4\\ & & \\ & & \end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}2x_1+3x_2+4x_3\\ \\ \end{bmatrix}$$

此矩阵说明，如何直接计算 $A\mathbf{x}$ 中的第一个元素，而不必像（7）中那样写出所有运算步骤. 类似地，$A\mathbf{x}$ 中的第二个元素可以直接把 A 的第二行各元素与 $\mathbf{x}$ 中对应元素相乘再求和得出：

$$\begin{bmatrix} & & \\ -1 & 5 & -3\\ & & \end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}\\ -x_1+5x_2-3x_3\\ \\ \end{bmatrix}$$

类似地，$A\mathbf{x}$ 中第 3 个元素可由 A 的第 3 行与 $\mathbf{x}$ 的元素算出. ■

> **计算 $A\mathbf{x}$ 的行——向量规则**
>
> 若乘积 $A\mathbf{x}$ 有定义，则 $A\mathbf{x}$ 中的第 i 个元素是 A 的第 i 行元素与 $\mathbf{x}$ 的相应元素乘积之和.

例 5

a. $\begin{bmatrix}1 & 2 & -1\\ 0 & -5 & 3\end{bmatrix}\begin{bmatrix}4\\ 3\\ 7\end{bmatrix}=\begin{bmatrix}1\cdot 4+2\cdot 3+(-1)\cdot 7\\ 0\cdot 4+(-5)\cdot 3+3\cdot 7\end{bmatrix}=\begin{bmatrix}3\\ 6\end{bmatrix}$

$$\text{b.}\ \begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 2\cdot 4+(-3)\cdot 7 \\ 8\cdot 4+0\cdot 7 \\ (-5)\cdot 4+2\cdot 7 \end{bmatrix} = \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}$$

$$\text{c.}\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r \\ s \\ t \end{bmatrix} = \begin{bmatrix} 1\cdot r+0\cdot s+0\cdot t \\ 0\cdot r+1\cdot s+0\cdot t \\ 0\cdot r+0\cdot s+1\cdot t \end{bmatrix} = \begin{bmatrix} r \\ s \\ t \end{bmatrix}$$

■

例5（c）中的矩阵，主对角线上元素为1，其他位置上元素为0，这个矩阵称为**单位矩阵**，并记为I.（c）中的计算说明，对任意$\mathbb{R}^3$中的$\boldsymbol{x}$，$I\boldsymbol{x}=\boldsymbol{x}$. 类似地，有$n\times n$单位矩阵，有时记为$I_n$，如（c）中所示，对任意$\mathbb{R}^n$中的$\boldsymbol{x}$，$I_n\boldsymbol{x}=\boldsymbol{x}$.

矩阵-向量积$A\boldsymbol{x}$的性质

下面定理中的事实是重要的，在本书里经常使用，它们的证明依赖于$A\boldsymbol{x}$的定义及$\mathbb{R}^n$的代数性质.

定理5 若A是$m\times n$矩阵，$\boldsymbol{u}$和$\boldsymbol{v}$是$\mathbb{R}^n$中向量，c是标量，则

a. $A(\boldsymbol{u}+\boldsymbol{v})=A\boldsymbol{u}+A\boldsymbol{v}$.

b. $A(c\boldsymbol{u})=c(A\boldsymbol{u})$.

证 为简单起见，取$n=3$，$A=[\boldsymbol{a}_1\ \boldsymbol{a}_2\ \ \boldsymbol{a}_3]$, $\boldsymbol{u},\boldsymbol{v}$为$\mathbb{R}^3$中的向量（一般情况的证明类似），对$i=1,2,3$，设$u_i$和$v_i$分别为$\boldsymbol{u}$和$\boldsymbol{v}$的第$i$个元素. 为证明（a），把$A(\boldsymbol{u}+\boldsymbol{v})$作为$A$的各列的以$\boldsymbol{u}+\boldsymbol{v}$的元素为权的线性组合来计算.

$$A(\boldsymbol{u}+\boldsymbol{v}) = [\boldsymbol{a}_1\ \ \boldsymbol{a}_2\ \ \boldsymbol{a}_3]\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \\ u_3+v_3 \end{bmatrix}$$

$\boldsymbol{u}+\boldsymbol{v}$ 的元素

$$= (u_1+v_1)\boldsymbol{a}_1 + (u_2+v_2)\boldsymbol{a}_2 + (u_3+v_3)\boldsymbol{a}_3$$

A 的各列

$$= (u_1\boldsymbol{a}_1+u_2\boldsymbol{a}_2+u_3\boldsymbol{a}_3) + (v_1\boldsymbol{a}_1+v_2\boldsymbol{a}_2+v_3\boldsymbol{a}_3)$$

$$= A\boldsymbol{u}+A\boldsymbol{v}$$

为证明（b），把$A(c\boldsymbol{u})$作为A的各列以$c\boldsymbol{u}$的各元素作为权的线性组合来计算.

$$A(c\boldsymbol{u}) = [\boldsymbol{a}_1\ \ \boldsymbol{a}_2\ \ \boldsymbol{a}_3]\begin{bmatrix} cu_1 \\ cu_2 \\ cu_3 \end{bmatrix} = (cu_1)\boldsymbol{a}_1+(cu_2)\boldsymbol{a}_2+(cu_3)\boldsymbol{a}_3$$

$$= c(u_1\boldsymbol{a}_1)+c(u_2\boldsymbol{a}_2)+c(u_3\boldsymbol{a}_3)$$

$$= c(u_1\boldsymbol{a}_1+u_2\boldsymbol{a}_2+u_3\boldsymbol{a}_3)$$

$$= c(A\boldsymbol{u})$$

■

数值计算的注解 为使计算$A\boldsymbol{x}$的计算机算法最优化，一系列的计算对存储在相连的存储单元中的数据进行. 矩阵计算中最广泛运用的算法是用Fortran语言写成的，它把矩阵作为若干列来存储，这样的算法把$A\boldsymbol{x}$作为A的列的线性组合来计算. 对比而言，

若程序用通用的 C 语言来写，它把矩阵按行存储，$A\boldsymbol{x}$ 就必须用另一种规则计算，这种算法使用 A 的行.

定理 4 的证明 如定理 4 后面所指出的，命题（a）、（b）和（c）逻辑上等价，因此，这里只需证明（对任意矩阵 A）命题（a）和（d）同时为真，或同时为假，就可以建立四个命题的等价性.

设 U 为 A 的阶梯形. 给定 $\mathbb{R}^m$ 中的 $\boldsymbol{b}$，我们可把增广矩阵 $[A\ \ \boldsymbol{b}]$ 行化简为增广矩阵 $[U\ \ \boldsymbol{d}]$，这里 $\boldsymbol{d}$ 是 $\mathbb{R}^m$ 中的某个向量.

$$[A\ \ \boldsymbol{b}] \sim \cdots \sim [U\ \ \boldsymbol{d}]$$

若（d）成立，U 的每一行包含一个主元位置而在增广列中不可能有主元. 故对任意 $\boldsymbol{b}$，$A\boldsymbol{x}=\boldsymbol{b}$ 有解，故（a）成立. 若（d）不成立，U 的最后一行都是 0，设 $\boldsymbol{d}$ 是最后一个元素为 1 的向量，于是 $[U\ \ \boldsymbol{d}]$ 代表一个不相容的方程组，因行变换是可逆的，$[U\ \ \boldsymbol{d}]$ 可变换为形如 $[A\ \ \boldsymbol{b}]$ 的矩阵，所得方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 也是不相容的，故（a）也不成立. ■

练习题

1. 设 $A=\begin{bmatrix}1&5&-2&0\\-3&1&9&-5\\4&-8&-1&7\end{bmatrix}$，$\boldsymbol{p}=\begin{bmatrix}3\\-2\\0\\-4\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}-7\\9\\0\end{bmatrix}$，可以证明 $\boldsymbol{p}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的一个解. 应用这个事实把 $\boldsymbol{b}$ 表示为 A 的列的线性组合.
2. 设 $A=\begin{bmatrix}2&5\\3&1\end{bmatrix}$，$\boldsymbol{u}=\begin{bmatrix}4\\-1\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}-3\\5\end{bmatrix}$，计算 $A(\boldsymbol{u}+\boldsymbol{v})$ 和 $A\boldsymbol{u}+A\boldsymbol{v}$ 来验证定理 5（a）.

习题 1.4

计算习题 1~4 中的乘积，(a) 像例 1 那样使用定义，(b) 使用计算 $A\boldsymbol{x}$ 的行向量规则. 若某一个乘积没有定义，加以说明.

1. $\begin{bmatrix}-4&2\\1&6\\0&1\end{bmatrix}\begin{bmatrix}3\\-2\\7\end{bmatrix}$

2. $\begin{bmatrix}1\\3\\1\end{bmatrix}\begin{bmatrix}5\\-1\end{bmatrix}$

3. $\begin{bmatrix}1&2\\-3&1\\1&6\end{bmatrix}\begin{bmatrix}-2\\3\end{bmatrix}$

4. $\begin{bmatrix}1&3&-4\\3&2&1\end{bmatrix}\begin{bmatrix}1\\2\\1\end{bmatrix}$

在习题 5~8 中，使用 $A\boldsymbol{x}$ 的定义把矩阵方程写成向量方程，或反过来.

5. $\begin{bmatrix}1&2&-3&1\\-2&-3&1&-1\end{bmatrix}\begin{bmatrix}2\\-1\\1\\-1\end{bmatrix}=\begin{bmatrix}-4\\1\end{bmatrix}$

6. $\begin{bmatrix}2&-3\\3&2\\8&-5\\-2&1\end{bmatrix}\begin{bmatrix}-3\\5\end{bmatrix}=\begin{bmatrix}-21\\1\\-49\\11\end{bmatrix}$

7. $x_1\begin{bmatrix}4\\-1\\7\\-4\end{bmatrix}+x_2\begin{bmatrix}-5\\3\\-5\\1\end{bmatrix}+x_3\begin{bmatrix}7\\-8\\0\\2\end{bmatrix}=\begin{bmatrix}6\\-8\\0\\-7\end{bmatrix}$

8. $z_1\begin{bmatrix}2\\-4\end{bmatrix}+z_2\begin{bmatrix}-1\\5\end{bmatrix}+z_3\begin{bmatrix}-4\\3\end{bmatrix}+z_4\begin{bmatrix}0\\2\end{bmatrix}=\begin{bmatrix}5\\12\end{bmatrix}$

在习题 9~10 中，将方程组写成向量方程和矩阵方程.

9. $\begin{aligned}5x_1+x_2-3x_3&=8\\2x_2+4x_3&=0\end{aligned}$

10. $\begin{aligned}4x_1-\ \ x_2&=8\\5x_1+3x_2&=2\\3x_1-\ \ x_2&=1\end{aligned}$

在习题 11~12 中，给定 A 和 $\boldsymbol{b}$，写出对应矩阵方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的增广矩阵并求解，将解表示成向量形式.

11. $A=\begin{bmatrix}1&3&-4\\1&5&2\\-3&-7&6\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}-2\\4\\12\end{bmatrix}$

12. $A=\begin{bmatrix}1&2&-1\\-3&-4&2\\5&2&3\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}1\\2\\-3\end{bmatrix}$

13. 设 $\boldsymbol{u}=\begin{bmatrix}0\\4\\4\end{bmatrix}$，$A=\begin{bmatrix}3&-5\\-2&6\\1&1\end{bmatrix}$，$\boldsymbol{u}$ 是否在由 A 的列所生成的 $\mathbb{R}^3$ 的子集中？（见图 1-20.）为什么？

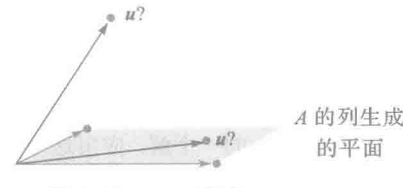


图 1-20　$\boldsymbol{u}$ 在何处

14. 设 $\boldsymbol{u}=\begin{bmatrix}4\\-1\\4\end{bmatrix}$，$A=\begin{bmatrix}2&5&-1\\0&1&-1\\1&2&0\end{bmatrix}$，$\boldsymbol{u}$ 是否在由 A 的列所生成的 $\mathbb{R}^3$ 的子集中？ 为什么？

15. 设 $A=\begin{bmatrix}3&-1\\-9&3\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}b_1\\b_2\end{bmatrix}$，证明方程 $A\boldsymbol{x}=\boldsymbol{b}$ 不是对一切 $\boldsymbol{b}$ 都相容，并说明使 $A\boldsymbol{x}=\boldsymbol{b}$ 相容的所有向量 $\boldsymbol{b}$ 的集合.

16. 又设 $A=\begin{bmatrix}1&-2&-1\\-2&2&0\\4&-1&3\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}$，重复 15 题.

习题 17~20 用到下面的矩阵 A 和 B，给出回答以及用到的定理.

$$A=\begin{bmatrix}1&3&0&3\\-1&-1&-1&1\\0&-4&2&-8\\2&0&3&-1\end{bmatrix}\quad B=\begin{bmatrix}1&4&1&2\\0&1&3&-4\\0&2&6&7\\2&9&5&-7\end{bmatrix}$$

17. A 中有多少行包含主元位置？方程 $A\boldsymbol{x}=\boldsymbol{b}$ 是否对 $\mathbb{R}^4$ 中的每个 $\boldsymbol{b}$ 都有解？

18. B 的列是否可以生成 $\mathbb{R}^4$？方程 $B\boldsymbol{x}=\boldsymbol{y}$ 是否对 $\mathbb{R}^4$ 中的每个 $\boldsymbol{y}$ 都有解？

19. $\mathbb{R}^4$ 中的每个向量都可以写成矩阵 A 的列的线性组合吗？A 的列是否可以生成 $\mathbb{R}^4$？

20. $\mathbb{R}^4$ 中的每个向量都可以写成矩阵 B 的列的线性组合吗？ B 的列是否可以生成 $\mathbb{R}^3$？

21. 设 $\boldsymbol{v}_1=\begin{bmatrix}1\\0\\-1\\0\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}0\\-1\\0\\1\end{bmatrix}$，$\boldsymbol{v}_3=\begin{bmatrix}1\\0\\0\\-1\end{bmatrix}$，$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 是否生成 $\mathbb{R}^4$？为什么？

22. 设 $\boldsymbol{v}_1=\begin{bmatrix}0\\0\\-3\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}0\\-3\\9\end{bmatrix}$，$\boldsymbol{v}_3=\begin{bmatrix}4\\-2\\-6\end{bmatrix}$，$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 是否生成 $\mathbb{R}^3$？

习题 23~24 中，判断各命题的真假，给出理由.

23. a. 方程 $A\boldsymbol{x}=\boldsymbol{b}$ 称为向量方程.
    b. 向量 $\boldsymbol{b}$ 是矩阵 A 的列的线性组合，当且仅当方程 $A\boldsymbol{x}=\boldsymbol{b}$ 至少有一个解.
    c. 方程 $A\boldsymbol{x}=\boldsymbol{b}$ 相容，当且仅当增广矩阵 $[A\ \ \boldsymbol{b}]$ 的每一行有一个主元位置.
    d. 乘积 $A\boldsymbol{x}$ 的第一个元素是乘积的和.
    e. 若 $m\times n$ 矩阵 A 的列生成 $\mathbb{R}^m$，则对 $\mathbb{R}^m$ 中任意的 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 相容.
    f. 若 A 是 $m\times n$ 矩阵，且方程 $A\boldsymbol{x}=\boldsymbol{b}$ 对 $\mathbb{R}^m$ 中某个 $\boldsymbol{b}$ 是不相容的，则 A 不能在每一行都有一个主元位置.

24. a. 每个矩阵方程 $A\boldsymbol{x}=\boldsymbol{b}$ 对应一个有相同解集的向量方程.
    b. 若方程，$A\boldsymbol{x}=\boldsymbol{b}$ 不相容，则 $\boldsymbol{b}$ 不属于 A 的列生成的集合.
    c. 向量的任何线性组合总可以写成 $A\boldsymbol{x}$ 的形式，A 是适当的矩阵，$\boldsymbol{x}$ 是适当的向量.
    d. 若增广矩阵 $[A\ \ \boldsymbol{b}]$ 在每一行有一个主元位置，则方程 $A\boldsymbol{x}=\boldsymbol{b}$ 不相容.
    e. 增广矩阵为 $[\boldsymbol{a}_1\ \ \boldsymbol{a}_2\ \ \boldsymbol{a}_3\ \ \boldsymbol{b}]$ 的线性方程组的解集与方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的解集相同，其中

$A=[\boldsymbol{a}_1\ \ \boldsymbol{a}_2\ \ \boldsymbol{a}_3]$.

f. 若 A 是 $m\times n$ 矩阵，它的列不生成 $\mathbb{R}^m$，则对 $\mathbb{R}^m$ 中某个 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 不相容.

25. 由等式 $\begin{bmatrix}4 & -3 & 1\\ 5 & -2 & 5\\ -6 & 2 & -3\end{bmatrix}\begin{bmatrix}-3\\ -1\\ 2\end{bmatrix}=\begin{bmatrix}-7\\ -3\\ 10\end{bmatrix}$ 求出标量 c_1,c_2,c_3（不用行变换）使

$$\begin{bmatrix}-7\\ -3\\ 10\end{bmatrix}=c_1\begin{bmatrix}4\\ 5\\ -6\end{bmatrix}+c_2\begin{bmatrix}-3\\ -2\\ 2\end{bmatrix}+c_3\begin{bmatrix}1\\ 5\\ -3\end{bmatrix}$$

26. 设 $\boldsymbol{u}=\begin{bmatrix}7\\ 2\\ 5\end{bmatrix}, \boldsymbol{v}=\begin{bmatrix}3\\ 1\\ 3\end{bmatrix}, \boldsymbol{w}=\begin{bmatrix}5\\ 1\\ 1\end{bmatrix}$，已知 $3\boldsymbol{u}-5\boldsymbol{v}-\boldsymbol{w}=\boldsymbol{0}$，

解方程（不用行变换）$\begin{bmatrix}7 & 3\\ 2 & 1\\ 5 & 3\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}5\\ 1\\ 1\end{bmatrix}$.

27. 使用符号 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots$ 表示向量，$c_1,c_2,\cdots$ 表示标量，重写下面的（数值）矩阵方程. 给出每个符号表示的意义. 矩阵方程如下：

$$\begin{bmatrix}-3 & 5 & -4 & 9 & 7\\ 5 & 8 & 1 & -2 & -4\end{bmatrix}\begin{bmatrix}-3\\ 1\\ 2\\ -1\\ 2\end{bmatrix}=\begin{bmatrix}11\\ -11\end{bmatrix}$$

28. 设 $\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3$ 和 $\boldsymbol{v}$ 是 $\mathbb{R}^5$ 中的向量，x_1,x_2,x_3 是标量，将向量方程 $x_1\boldsymbol{q}_1+x_2\boldsymbol{q}_2+x_3\boldsymbol{q}_3=\boldsymbol{v}$ 写成一个矩阵方程，注意你选用的符号.

29. 构造一个 3×3 非阶梯形矩阵，使得矩阵的列可以生成 $\mathbb{R}^3$. 说明你构造的矩阵具有这种属性.

30. 构造一个 3×3 非阶梯形矩阵，使得矩阵的列不可以生成 $\mathbb{R}^3$. 说明你构造的矩阵具有这种属性.

31. 设 A 是 3×2 矩阵，说明为什么方程 $A\boldsymbol{x}=\boldsymbol{b}$ 不可能对所有 $\mathbb{R}^3$ 中的向量 $\boldsymbol{b}$ 都是相容的. 推广你的结论到任意行数多于列数的矩阵 A.

32. $\mathbb{R}^4$ 中的 3 个向量能否生成整个 $\mathbb{R}^4$？说明理由. 当 $n>m$ 时，$\mathbb{R}^m$ 中的 n 个向量是能否生成 $\mathbb{R}^m$？

33. 设 A 是 4×3 矩阵，$\boldsymbol{b}$ 是 $\mathbb{R}^4$ 中的任一向量，且 $A\boldsymbol{x}=\boldsymbol{b}$ 有唯一解. 由此可知 A 的简化阶梯形是怎样的？给出理由.

34. 设 A 是 3×4 矩阵，$\boldsymbol{v}_1,\boldsymbol{v}_2$ 为 $\mathbb{R}^3$ 中的向量，$\boldsymbol{w}=\boldsymbol{v}_1+\boldsymbol{v}_2$，设对 $\mathbb{R}^4$ 中的向量 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$，$\boldsymbol{v}_1=A\boldsymbol{u}_1$，$\boldsymbol{v}_2=A\boldsymbol{u}_2$. 为什么方程 $A\boldsymbol{x}=\boldsymbol{w}$ 相容？（注：$\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 是向量而不是向量中的数值元素.）

35. 设 A 是 5×3 矩阵，$\boldsymbol{y}$ 是 $\mathbb{R}^3$ 中的向量，$\boldsymbol{z}$ 是 $\mathbb{R}^5$ 中的向量. 又设 $A\boldsymbol{y}=\boldsymbol{z}$，什么事实使你断定方程 $A\boldsymbol{x}=5\boldsymbol{z}$ 是相容的？

36. 设 A 是 4×4 矩阵，$\boldsymbol{b}$ 是 $\mathbb{R}^3$ 中的任一向量，且 $A\boldsymbol{x}=\boldsymbol{b}$ 有唯一解. 说明为什么 A 的列一定可以生成 $\mathbb{R}^4$？

[M]习题 37~40 中，确定矩阵各列能否生成 $\mathbb{R}^4$.

37. $\begin{bmatrix}7 & 2 & -5 & 8\\ -5 & -3 & 4 & -9\\ 6 & 10 & -2 & 7\\ -7 & 9 & 2 & 15\end{bmatrix}$

38. $\begin{bmatrix}4 & -5 & -1 & 8\\ 3 & -7 & -4 & 2\\ 5 & -6 & -1 & 4\\ 9 & 1 & 10 & 7\end{bmatrix}$

39. $\begin{bmatrix}10 & -7 & 1 & 4 & 6\\ -8 & 4 & -6 & -10 & -3\\ -7 & 11 & -5 & -1 & -8\\ 3 & -1 & 10 & 12 & 12\end{bmatrix}$

40. $\begin{bmatrix}5 & 11 & -6 & -7 & 12\\ -7 & -3 & -4 & 6 & -9\\ 11 & 5 & 6 & -9 & -3\\ -3 & 4 & -7 & 2 & 7\end{bmatrix}$

41. [M]在习题 39 中去掉矩阵的某一列，使剩下的各列仍然可以生成 $\mathbb{R}^4$.

42. [M]在习题 40 中去掉矩阵的某一列，使剩下的各列仍然可以生成 $\mathbb{R}^4$. 能否去掉更多的列？

练习题答案

1. 矩阵方程 $\begin{bmatrix} 1 & 5 & -2 & 0 \\ -3 & 1 & 9 & -5 \\ 4 & -8 & -1 & 7 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \\ 0 \\ -4 \end{bmatrix}=\begin{bmatrix} -7 \\ 9 \\ 0 \end{bmatrix}$等价于向量方程$3\begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix}-2\begin{bmatrix} 5 \\ 1 \\ -8 \end{bmatrix}+0\begin{bmatrix} -2 \\ 9 \\ -1 \end{bmatrix}-4\begin{bmatrix} 0 \\ -5 \\ 7 \end{bmatrix}=\begin{bmatrix} -7 \\ 9 \\ 0 \end{bmatrix}$，它表示 $\boldsymbol{b}$ 是 A 的各列的线性组合.

2. $\boldsymbol{u}+\boldsymbol{v}=\begin{bmatrix} 4 \\ -1 \end{bmatrix}+\begin{bmatrix} -3 \\ 5 \end{bmatrix}=\begin{bmatrix} 1 \\ 4 \end{bmatrix}$

$$A(\boldsymbol{u}+\boldsymbol{v})=\begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 4 \end{bmatrix}=\begin{bmatrix} 2+20 \\ 3+4 \end{bmatrix}=\begin{bmatrix} 22 \\ 7 \end{bmatrix}$$

$$A\boldsymbol{u}+A\boldsymbol{v}=\begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ -1 \end{bmatrix}+\begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} -3 \\ 5 \end{bmatrix}=\begin{bmatrix} 3 \\ 11 \end{bmatrix}+\begin{bmatrix} 19 \\ -4 \end{bmatrix}=\begin{bmatrix} 22 \\ 7 \end{bmatrix}$$

1.5 线性方程组的解集

线性方程组的解集是线性代数研究的重要对象，它们出现在许多不同的问题中. 本节使用向量符号给出这样的解集的显式表示以及几何解释.

齐次线性方程组

线性方程组称为**齐次的**，若它可写成 $A\boldsymbol{x}=\boldsymbol{0}$ 的形式，其中 A 是 $m\times n$ 矩阵而 $\boldsymbol{0}$ 是 $\mathbb{R}^m$ 中的零向量. 这样的方程组至少有一个解. 即 $\boldsymbol{x}=\boldsymbol{0}$ ($\mathbb{R}^n$ 中的零向量)，这个解称为它的**平凡解**. 对给定方程 $A\boldsymbol{x}=\boldsymbol{0}$，重要的是它是否有**非平凡解**，即满足 $A\boldsymbol{x}=\boldsymbol{0}$ 的非零向量 $\boldsymbol{x}$. 由 1.2 节解的存在性与唯一性定理（定理 2），得出以下事实.

> 齐次方程 $A\boldsymbol{x}=\boldsymbol{0}$ 有非平凡解，当且仅当方程至少有一个自由变量.

例 1 确定下列齐次方程组是否有非平凡解，并描述它的解集.

$$\begin{aligned} 3x_1+5x_2-4x_3&=0 \\ -3x_1-2x_2+4x_3&=0 \\ 6x_1+\ \ x_2-8x_3&=0 \end{aligned}$$

解 令 A 为该方程组的系数矩阵，用行化简法把增广矩阵$[A\ \ \boldsymbol{0}]$化为阶梯形.

$$\begin{bmatrix} 3 & 5 & -4 & 0 \\ -3 & -2 & 4 & 0 \\ 6 & 1 & -8 & 0 \end{bmatrix}\sim\begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & -9 & 0 & 0 \end{bmatrix}\sim\begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

因 x_3 是自由变量，$A\boldsymbol{x}=\boldsymbol{0}$ 有非平凡解（对 x_3 的每一选择都有一个解），为描述解集，继续把$[A\ \ \boldsymbol{0}]$化为简化阶梯形：

$$\begin{bmatrix} 1 & 0 & -\frac{4}{3} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{aligned} x_1 \quad -\frac{4}{3}x_3 &= 0 \\ x_2 \qquad\quad &= 0 \\ 0 &= 0 \end{aligned}$$

解出基本变量 x_1 和 x_2 得 $x_1 = \frac{4}{3}x_3, x_2 = 0, x_3$ 是自由变量，$A\boldsymbol{x} = \boldsymbol{0}$ 的通解有向量形式

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix} = x_3\boldsymbol{v}, \text{ 其中 } \boldsymbol{v} = \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}$$

这里 x_3 由通解向量的表达式中作为公因子提出来. 这说明本例中 $A\boldsymbol{x} = \boldsymbol{0}$ 的每一个解都是 $\boldsymbol{v}$ 的倍数. 平凡解可由 $x_3 = 0$ 得到. 几何意义下，解集是 $\mathbb{R}^3$ 中通过 $\boldsymbol{0}$ 的直线，见图 1-21.

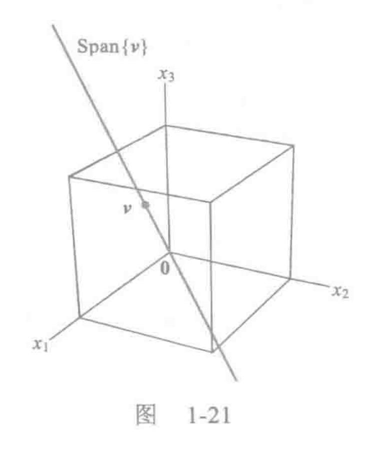


图 1-21 ■

注意，非平凡解向量 $\boldsymbol{x}$ 可能有些零元素，只要不是所有元素都是 0 就行.

例 2 单一方程也可看作方程组，描述下列齐次"方程组"的解集.

$$10x_1 - 3x_2 - 2x_3 = 0 \tag{1}$$

解 这里无须矩阵记号. 用自由变量表示基本变量 x_1. 通解为

$$x_1 = 0.3x_2 + 0.2x_3$$

x_2 和 x_3 为自由变量. 写成向量形式，通解为

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.3x_2 + 0.2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.3x_2 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.2x_3 \\ 0 \\ x_3 \end{bmatrix}$$

$$= x_2 \underset{\uparrow \atop \boldsymbol{u}}{\begin{bmatrix} 0.3 \\ 1 \\ 0 \end{bmatrix}} + x_3 \underset{\uparrow \atop \boldsymbol{v}}{\begin{bmatrix} 0.2 \\ 0 \\ 1 \end{bmatrix}} \quad (x_2, x_3 \text{ 为自由变量}) \tag{2}$$

计算表明，方程（1）的每个解都是向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的线性组合，如（2）式所示. 即解集为 Span$\{\boldsymbol{u},\boldsymbol{v}\}$，因为，$\boldsymbol{u}$ 不是 $\boldsymbol{v}$ 的倍数，解集是通过原点的一个平面. 见图 1-22.

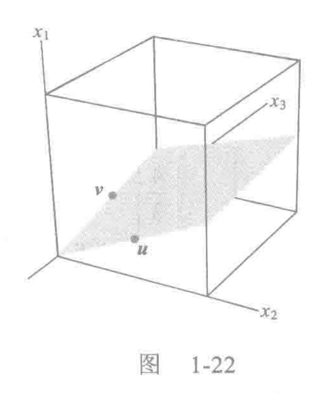


图 1-22 ■

例 1 和例 2 以及后面的练习，说明齐次方程 $A\boldsymbol{x}=\boldsymbol{0}$ 总可表示为 Span$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$，其中 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 是适当的解向量. 若唯一解是零向量，则解集就是 Span$\{\boldsymbol{0}\}$，若方程 $A\boldsymbol{x}=\boldsymbol{0}$ 仅有一个自由变量，解集是通过原点的一条直线，见图 1-21. 若有两个或更多自由变量，那么图 1-22 的通过原点的平面就给出 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集的一个很好的图形说明. 注意，类似的图可用来解释 Span$\{\boldsymbol{u},\boldsymbol{v}\}$，即使 $\boldsymbol{u},\boldsymbol{v}$ 并不是 $A\boldsymbol{x}=\boldsymbol{0}$ 的解，见 1.3 节图 1-17.

参数向量形式

最初的方程（1）是例 2 中的平面的隐式描述，解此方程就是要找这个平面的显式描述，就是说，将它作为 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 所生成的子集. 方程（2）称为平面的**参数向量方程**. 有时也可写为

$$\boldsymbol{x}=s\boldsymbol{u}+t\boldsymbol{v} \qquad (s,t \text{ 为实数})$$

来强调参数可取任何实数值，例 1 中，方程 $\boldsymbol{x}=x_3\boldsymbol{v}$ (x_3 是自由变量)，或 $\boldsymbol{x}=t\boldsymbol{v}$ (t 为实数)，是直线的参数方程，当解集用向量显式表示为如例 1 和例 2 时，我们称之为解的**参数向量形式**.

非齐次方程组的解

当非齐次线性方程组有许多解时，一般可表示为参数向量形式，即由一个向量加上满足对应的齐次方程的一些向量的任意线性组合的形式.

例 3 描述 $A\boldsymbol{x}=\boldsymbol{b}$ 的解，其中

$$A=\begin{bmatrix}3 & 5 & -4\\ -3 & -2 & 4\\ 6 & 1 & -8\end{bmatrix},\quad \boldsymbol{b}=\begin{bmatrix}7\\ -1\\ -4\end{bmatrix}$$

解 这里 A 就是例 1 的系数矩阵. 对 $[A\ \ \boldsymbol{b}]$ 作行变换得

$$\begin{bmatrix}3 & 5 & -4 & 7\\ -3 & -2 & 4 & -1\\ 6 & 1 & -8 & -4\end{bmatrix}\sim\begin{bmatrix}1 & 0 & -\frac{4}{3} & -1\\ 0 & 1 & 0 & 2\\ 0 & 0 & 0 & 0\end{bmatrix} \qquad \begin{aligned}x_1 \quad -\tfrac{4}{3}x_3 &= -1\\ x_2 \quad &= 2\\ 0 &= 0\end{aligned}$$

所以 $x_1=-1+\frac{4}{3}x_3, x_2=2, x_3$ 为自由变量，$A\boldsymbol{x}=\boldsymbol{b}$ 的通解可写成向量形式

$$\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}-1+\frac{4}{3}x_3\\2\\x_3\end{bmatrix}=\begin{bmatrix}-1\\2\\0\end{bmatrix}+\begin{bmatrix}\frac{4}{3}x_3\\0\\x_3\end{bmatrix}=\underset{\uparrow\atop \boldsymbol{p}}{\begin{bmatrix}-1\\2\\0\end{bmatrix}}+x_3\underset{\uparrow\atop \boldsymbol{v}}{\begin{bmatrix}\frac{4}{3}\\0\\1\end{bmatrix}}$$

方程 $\boldsymbol{x}=\boldsymbol{p}+x_3\boldsymbol{v}$，或用 t 表示一般参数，

$$\boldsymbol{x}=\boldsymbol{p}+t\boldsymbol{v}\quad（t\text{ 为实数}）\tag{3}$$

就是用参数向量形式表示 $A\boldsymbol{x}=\boldsymbol{b}$ 的解集. 回忆例 1 中 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集有参数向量形式

$$\boldsymbol{x}=t\boldsymbol{v}\quad（t\text{ 为实数}）\tag{4}$$

（$\boldsymbol{v}$ 与（3）中的 $\boldsymbol{v}$ 相同），故 $A\boldsymbol{x}=\boldsymbol{b}$ 的解可由向量 $\boldsymbol{p}$ 加上 $A\boldsymbol{x}=\boldsymbol{0}$ 的解得到，向量 $\boldsymbol{p}$ 本身也是 $A\boldsymbol{x}=\boldsymbol{b}$ 的一个特解（在（3）中对应 $t=0$）. ■

为了从几何上描述 $A\boldsymbol{x}=\boldsymbol{b}$ 的解集，我们可以把向量加法解释为平移，给定 $\mathbb{R}^2$ 或 $\mathbb{R}^3$ 中的向量 $\boldsymbol{v}$ 与 $\boldsymbol{p}$，把 $\boldsymbol{p}$ 加上 $\boldsymbol{v}$ 的结果就是把 $\boldsymbol{v}$ 沿着平行于通过 $\boldsymbol{p}$ 与 $\boldsymbol{0}$ 的直线移动，我们称 $\boldsymbol{v}$ 被**平移** $\boldsymbol{p}$ 到 $\boldsymbol{v}+\boldsymbol{p}$，见图 1-23. 若 $\mathbb{R}^2$ 或 $\mathbb{R}^3$ 中的直线 L 上的每一点被平移 $\boldsymbol{p}$，就得到一条平行于 L 的直线，见图 1-24.

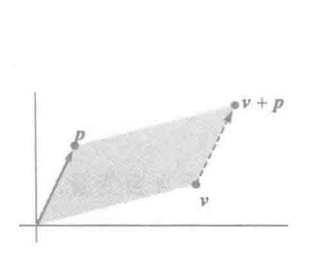


图 1-23　$\boldsymbol{v}$ 加上 $\boldsymbol{p}$，使 $\boldsymbol{v}$ 平移到 $\boldsymbol{v}+\boldsymbol{p}$

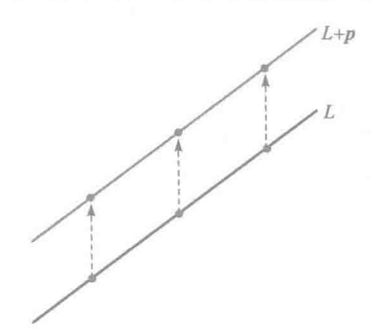


图 1-24　直线的平移

设 L 是通过 $\boldsymbol{0}$ 与 $\boldsymbol{v}$ 的直线，由方程（4）表示. L 的每个点加上 $\boldsymbol{p}$ 得到由方程（3）表示的平移后的直线. 注意 $\boldsymbol{p}$ 也在直线（3）上. 称（3）为**通过 $\boldsymbol{p}$ 平行于 $\boldsymbol{v}$ 的直线方程**. 于是 $A\boldsymbol{x}=\boldsymbol{b}$ 的解集是一条通过 $\boldsymbol{p}$ 而平行于 $A\boldsymbol{x}=\boldsymbol{0}$ 的直线. 图 1-25 说明了这一结论.

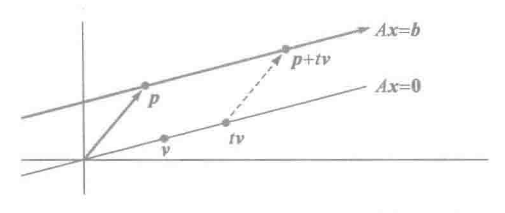


图 1-25　$A\boldsymbol{x}=\boldsymbol{b}$ 与 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集平行

图 1-25 中 $A\boldsymbol{x}=\boldsymbol{b}$ 和 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集之间的关系可以推广到任意相容的方程 $A\boldsymbol{x}=\boldsymbol{b}$，虽然当自由变量有多个时，解集将多于一条直线. 下列定理给出了这一结论，证明见习题 25.

定理 6 设方程 $A\boldsymbol{x}=\boldsymbol{b}$ 对某个 $\boldsymbol{b}$ 是相容的，$\boldsymbol{p}$ 为一个特解，则 $A\boldsymbol{x}=\boldsymbol{b}$ 的解集是所有形如 $\boldsymbol{w}=\boldsymbol{p}+\boldsymbol{v}_h$ 的向量的集，其中 $\boldsymbol{v}_h$ 是齐次方程 $A\boldsymbol{x}=\boldsymbol{0}$ 的任意一个解.

定理 6 说明若 $A\boldsymbol{x}=\boldsymbol{b}$ 有解，则解集可由 $A\boldsymbol{x}=\boldsymbol{0}$ 的解平移向量 $\boldsymbol{p}$ 得到，$\boldsymbol{p}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的任意一个特解，图 1-26 说明当有两个自由变量时的情形. 即使当 $n>3$ 时，相容方程组 $A\boldsymbol{x}=\boldsymbol{b}$（$\boldsymbol{b}\neq\boldsymbol{0}$）的解集也可想象为一个非零点或一条不通过原点的线或平面.

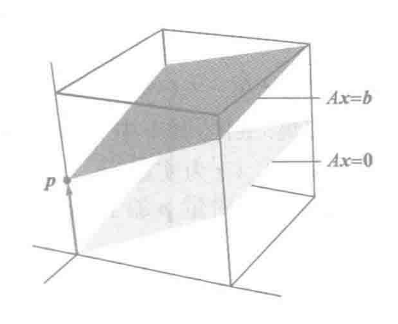


图 1-26 $A\boldsymbol{x}=\boldsymbol{b}$ 与 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集平行

警告 定理 6 与图 1-26 仅适用于方程 $A\boldsymbol{x}=\boldsymbol{b}$ 至少有一个非零解 $\boldsymbol{p}$ 的前提下.

下列算法总结了例 1、2 和 3 中的计算.

> **把相容方程组的解集表示成参数向量形式**
>
> 1. 把增广矩阵行化简为简化阶梯形.
> 2. 把每个基本变量用自由变量表示.
> 3. 把一般解 $\boldsymbol{x}$ 表示成向量，如果有自由变量，其元素依赖于自由变量.
> 4. 把 $\boldsymbol{x}$ 分解为向量（元素为常数）的线性组合，用自由变量作为参数.

练习题

1. 下列两个方程都确定 $\mathbb{R}^3$ 中的平面，它们是否相交？如果相交的话，描述它们的交集.

$$\begin{aligned} x_1+4x_2-5x_3&=0\\ 2x_1-\ \ x_2+8x_3&=9 \end{aligned}$$

2. 写出方程 $10x_1-3x_2-2x_3=7$ 的参数向量形式的通解，讨论这个解集与例 2 中的解集的关系.

习题 1.5

在习题 1~4 中，确定方程组是否有非平凡解，使用尽可能少的行运算.

1. $\begin{aligned} 2x_1-5x_2+8x_3&=0\\ -2x_1-7x_2+\ \ x_3&=0\\ 4x_1+2x_2+7x_3&=0 \end{aligned}$

2. $\begin{aligned} x_1-2x_2+3x_3&=0\\ -2x_1-3x_2-4x_3&=0\\ 2x_1-4x_2+9x_3&=0 \end{aligned}$

3. $\begin{aligned} -3x_1+4x_2-8x_3&=0\\ -2x_1+5x_2+4x_3&=0 \end{aligned}$

4. $\begin{aligned} 5x_1-3x_2+2x_3&=0\\ -3x_1-4x_2+2x_3&=0 \end{aligned}$

在习题 5、6 中，用例 1、2 的方法把给出的各线性方程组的解集用参数向量形式表示出来.

5. $\begin{aligned} x_1+3x_2+\ \ x_3&=0\\ -4x_1-9x_2+2x_3&=0\\ -3x_2-6x_3&=0 \end{aligned}$

6. $\begin{aligned} x_1+3x_2-5x_3&=0\\ x_1+4x_2-8x_3&=0\\ -3x_1-7x_2+9x_3&=0 \end{aligned}$

在习题 7~12 中，把方程 $A\boldsymbol{x}=\boldsymbol{0}$ 的解用参数向量形式表示出来，其中 A 行等价于给定的矩阵.

7. $\begin{bmatrix}1 & 3 & -3 & 7\\0 & 1 & -4 & 5\end{bmatrix}$

8. $\begin{bmatrix}1 & -3 & -8 & 5\\0 & 1 & 2 & -4\end{bmatrix}$

9. $\begin{bmatrix}3 & -6 & 6\\-2 & 4 & -2\end{bmatrix}$

10. $\begin{bmatrix}-1 & -4 & 0 & -4\\2 & -8 & 0 & 8\end{bmatrix}$

11. $\begin{bmatrix}1 & -4 & -2 & 0 & 3 & -5\\0 & 0 & 1 & 0 & 0 & -1\\0 & 0 & 0 & 0 & 1 & -4\\0 & 0 & 0 & 0 & 0 & 0\end{bmatrix}$

12. $\begin{bmatrix}1 & -2 & 3 & -6 & 5 & 0\\0 & 0 & 1 & 1 & 4 & -6\\0 & 0 & 0 & 0 & 0 & 1\\0 & 0 & 0 & 0 & 0 & 0\end{bmatrix}$

13. 设某方程组的解集表示为 $x_1 = 5+4x_3, x_2 = -2-7x_3, x_3$ 为自由变量. 用向量把它表示成 $\mathbb{R}^3$ 中的直线.

14. 设某方程组的解集表示为 $x_1 = 5x_4, x_2 = 3-2x_4, x_3 = 2+5x_4, x_4$ 为自由变量. 用向量把它表示成 $\mathbb{R}^4$ 中的直线.

15. 说明和比较 $x_1+5x_2-3x_3=0$ 与 $x_1+5x_2-3x_3=-2$ 的解集.

16. 说明和比较 $x_1-2x_2+3x_3=0$ 与 $x_1-2x_2+3x_3=4$ 的解集.

17. 用例 3 的方法以参数向量形式表示下列方程组的解. 给出该解集的几何解释并与习题 5 的解集做比较.

$$\begin{aligned}2x_1+2x_2+4x_3 &= 8\\-4x_1-4x_2-8x_3 &= -16\\-3x_2-3x_3 &= 12\end{aligned}$$

18. 和习题 17 一样，把下列方程组的解集表示出来，给出几何解释并与习题 6 的解集做比较.

$$\begin{aligned}x_1+2x_2-3x_3 &= 5\\2x_1+x_2-3x_3 &= 13\\-x_1+x_2 \quad\quad &= -8\end{aligned}$$

在习题 19~20 中，求出通过 $\boldsymbol{a}$ 且平行于 $\boldsymbol{b}$ 的直线的参数方程.

19. $\boldsymbol{a}=\begin{bmatrix}-2\\0\end{bmatrix}, \boldsymbol{b}=\begin{bmatrix}-5\\3\end{bmatrix}$

20. $\boldsymbol{a}=\begin{bmatrix}3\\-2\end{bmatrix}, \boldsymbol{b}=\begin{bmatrix}-7\\6\end{bmatrix}$

在习题 21~22 中，求出通过 $\boldsymbol{p}$ 与 $\boldsymbol{q}$ 的直线 M 的方程.（提示：M 平行于向量 $\boldsymbol{q}-\boldsymbol{p}$. 见图 1-27.）

21. $\boldsymbol{p}=\begin{bmatrix}3\\-3\end{bmatrix}, \boldsymbol{q}=\begin{bmatrix}4\\1\end{bmatrix}$

22. $\boldsymbol{p}=\begin{bmatrix}-3\\2\end{bmatrix}, \boldsymbol{q}=\begin{bmatrix}0\\-3\end{bmatrix}$

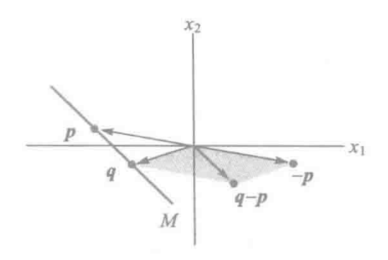


图 1-27 通过 $\boldsymbol{p}$ 与 $\boldsymbol{q}$ 的直线

在习题 23 和 24 中，判断每个命题的真假，并给出理由.

23. a. 齐次方程总是相容的.
   b. 方程 $A\boldsymbol{x}=\boldsymbol{0}$ 给出它的解集的显式表达式.
   c. 齐次方程 $A\boldsymbol{x}=\boldsymbol{0}$ 当且仅当方程至少有一个自由变量时有平凡解.
   d. 方程 $\boldsymbol{x}=\boldsymbol{p}+t\boldsymbol{v}$ 描述了一条直线，它通过 $\boldsymbol{v}$ 而平行于 $\boldsymbol{p}$.
   e. 方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的解集是所有形如 $\boldsymbol{w}=\boldsymbol{p}+\boldsymbol{v}_h$ 的向量的集，其中 $\boldsymbol{v}_h$ 是方程 $A\boldsymbol{x}=\boldsymbol{0}$ 的任意一解.

24. a. 齐次方程组可能不相容.
   b. 若 $\boldsymbol{x}$ 是 $A\boldsymbol{x}=\boldsymbol{0}$ 的非平凡解，则 $\boldsymbol{x}$ 的每个元素不等于零.
   c. 把一个向量加上 $\boldsymbol{p}$ 就是把该向量沿平行于 $\boldsymbol{p}$ 的方向移动.
   d. 方程 $A\boldsymbol{x}=\boldsymbol{b}$ 是齐次的，当且仅当零向量是它的解.
   e. 若 $A\boldsymbol{x}=\boldsymbol{b}$ 相容，则方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的解集可由平移 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集得到.

25. 证明定理 6：
   a. 设 $\boldsymbol{p}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的解，即 $A\boldsymbol{p}=\boldsymbol{b}$，设 $\boldsymbol{v}_h$ 为齐次方程 $A\boldsymbol{x}=\boldsymbol{0}$ 的解，$\boldsymbol{w}=\boldsymbol{p}+\boldsymbol{v}_h$. 证明 $\boldsymbol{w}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的解.

b. 设 $\boldsymbol{w}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的任意解，定义 $\boldsymbol{v}_h=\boldsymbol{w}-\boldsymbol{p}$，证明 $\boldsymbol{v}_h$ 是 $A\boldsymbol{x}=\boldsymbol{0}$ 的解．这说明，$A\boldsymbol{x}=\boldsymbol{b}$ 的任意解有形式 $\boldsymbol{w}=\boldsymbol{p}+\boldsymbol{v}_h$，$\boldsymbol{p}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的特解，$\boldsymbol{v}_h$ 是 $A\boldsymbol{x}=\boldsymbol{0}$ 的解．

26. 设 A 是 3×3 零矩阵（所有元素都是零），求方程 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集．

27. 设 $A\boldsymbol{x}=\boldsymbol{b}$ 有解，说明为什么当 $A\boldsymbol{x}=\boldsymbol{0}$ 仅有平凡解时，$A\boldsymbol{x}=\boldsymbol{b}$ 的解是唯一的．

在习题 28~31 中，（a）方程 $A\boldsymbol{x}=\boldsymbol{0}$ 是否有非平凡解，（b）方程 $A\boldsymbol{x}=\boldsymbol{b}$ 是否对每个 $\boldsymbol{b}$ 都至少有一个解？

28. A 是 3×3 矩阵，有 3 个主元位置．

29. A 是 4×4 矩阵，有 3 个主元位置．

30. A 是 2×5 矩阵，有 2 个主元位置．

31. A 是 3×2 矩阵，有 2 个主元位置．

32. 若 $\boldsymbol{b}\neq\boldsymbol{0}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的解集是否可能是通过原点的平面？说明理由．

33. 构造一 3×3 非零矩阵 A，使向量 $\begin{bmatrix}1\\1\\1\end{bmatrix}$ 是 $A\boldsymbol{x}=\boldsymbol{0}$ 的一个解．

34. 构造一 3×3 非零矩阵 A，使向量 $\begin{bmatrix}2\\-1\\1\end{bmatrix}$ 是 $A\boldsymbol{x}=\boldsymbol{0}$ 的一个解．

35. 给定 $A=\begin{bmatrix}-1&-3\\7&21\\-2&-6\end{bmatrix}$，用观察法求 $A\boldsymbol{x}=\boldsymbol{0}$ 的一个非平凡解．（提示：把方程 $A\boldsymbol{x}=\boldsymbol{0}$ 写成向量方程形式．）

36. 给定 $A=\begin{bmatrix}3&-2\\-6&4\\12&-8\end{bmatrix}$，用观察法求 $A\boldsymbol{x}=\boldsymbol{0}$ 的一个非平凡解．

37. 构造一 2×2 矩阵 A，使方程 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集是一条经点（4, 1）和原点的 $\mathbb{R}^2$ 中的直线．随后，在 $\mathbb{R}^2$ 中找一向量 $\boldsymbol{b}$ 使 $A\boldsymbol{x}=\boldsymbol{b}$ 的解集不是 $\mathbb{R}^2$ 中平行于 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集的直线．为什么这与定理 6 没有矛盾？

38. 设 A 是 $m\times n$ 矩阵，$\boldsymbol{w}$ 是 $\mathbb{R}^n$ 中满足 $A\boldsymbol{x}=\boldsymbol{0}$ 的向量．证明对任一数 c，向量 $c\boldsymbol{w}$ 也满足 $A\boldsymbol{x}=\boldsymbol{0}$．（即证明 $A(c\boldsymbol{w})=\boldsymbol{0}$．）

39. 设 A 是 $m\times n$ 矩阵，$\boldsymbol{v}$和$\boldsymbol{w}$ 是 $\mathbb{R}^n$ 中满足 $A\boldsymbol{v}=\boldsymbol{0}$ 和 $A\boldsymbol{w}=\boldsymbol{0}$ 的向量，解释为什么 $A(\boldsymbol{v}+\boldsymbol{w})$ 一定是零向量，以及对每一对标量 c 和 d，为什么 $A(c\boldsymbol{v}+d\boldsymbol{w})=\boldsymbol{0}$．

40. 设 A 是 3×3 矩阵，$\boldsymbol{y}$ 是 $\mathbb{R}^3$ 中的一个向量，且方程 $A\boldsymbol{x}=\boldsymbol{y}$ 无解，讨论是否存在 $\mathbb{R}^3$ 中的一个向量 $\boldsymbol{z}$，使方程 $A\boldsymbol{x}=\boldsymbol{z}$ 有唯一解？

练习题答案

1. 行化简增广矩阵

$$\begin{bmatrix}1&4&-5&0\\2&-1&8&9\end{bmatrix}\sim\begin{bmatrix}1&4&-5&0\\0&-9&18&9\end{bmatrix}\sim\begin{bmatrix}1&0&3&4\\0&1&-2&-1\end{bmatrix}$$

$$\begin{aligned}x_1\qquad+3x_3&=4\\x_2-2x_3&=-1\end{aligned}$$

因此 $x_1=4-3x_3, x_2=-1+2x_3, x_3$ 为自由变量．通解的向量形式为

$$\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}4-3x_3\\-1+2x_3\\x_3\end{bmatrix}=\underset{\substack{\uparrow\\ \boldsymbol{p}}}{\begin{bmatrix}4\\-1\\0\end{bmatrix}}+x_3\underset{\substack{\uparrow\\ \boldsymbol{v}}}{\begin{bmatrix}-3\\2\\1\end{bmatrix}}$$

两个平面的交是通过 $\boldsymbol{p}$ 平行于 $\boldsymbol{v}$ 的直线．

2. 增广矩阵$[10 \quad -3 \quad -2 \quad 7]$行等价于$[1 \quad -0.3 \quad -0.2 \quad 0.7]$，故通解为$x_1 = 0.7 + 0.3x_2 + 0.2x_3$，$x_2$和$x_3$是自由变量，即

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.7 + 0.3x_2 + 0.2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.7 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0.3 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0.2 \\ 0 \\ 1 \end{bmatrix}$$

$$= \boldsymbol{p} + x_2\boldsymbol{u} + x_3\boldsymbol{v}$$

非齐次方程$A\boldsymbol{x} = \boldsymbol{b}$的解集是平移过的平面$\boldsymbol{p} + \text{Span}\{\boldsymbol{u}, \boldsymbol{v}\}$，它经过$\boldsymbol{p}$且平行于例 2 中的齐次方程的解集.

1.6 线性方程组的应用

你也许希望现实生活中涉及线性代数的问题会是只有唯一解，或者可能无解. 本节的意图是要说明有多解的线性方程组是如何自然产生的. 这里的实例来自经济学、化学和网络流.

经济学中的齐次线性方程组

本章介绍中提到的 500 个变量的 500 个方程组成的方程组，现称为列昂惕夫“投入-产出”（或“生产”）模型.[㊀] 2.6 节将详细讨论这个模型，那时我们有更多的理论和更好的符号. 目前，我们先看一个简单的“交易模型”，这个模型也是由列昂惕夫提出的.

假设一个国家的经济可以划分为许多部门，如各种制造、交通、娱乐和服务业. 假设我们知道每个部门年度的总产出，并精确知道该总产出是如何在其他经济部门进行分配或“交易”的. 称一个部门产出的总货币价值为该产出的**价格**. 列昂惕夫证明了下面的结论.

存在能够指派给各个部门总产出的平衡价格，使得每个部门的总收入恰等于它的总支出.

下面的例子说明如何求平衡价格.

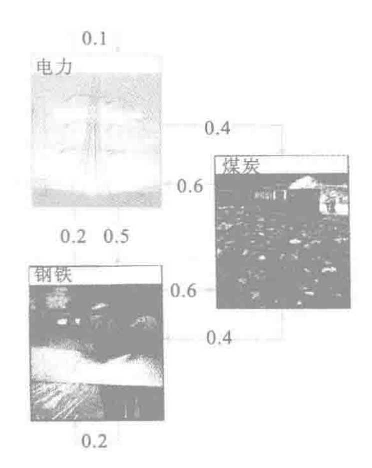
例 1 假设一个经济由煤炭、电力（电源）和钢铁三个部门组成，各部门之间的分配如表 1-1 所示，其中每一列中的数表示该部门总产出的比例. 如表 1-1 的第二列，将电力的总产出分配如下：40%给煤炭部门，50%给钢铁部门，剩下 10%给电力部门.（电力部门把这 10%作为运转费用.）因所有产出都必须分配，每一列的分数之和等于 1.

记符号p_C, p_E, p_S分别表示煤炭、电力和钢铁部门年度总产出的价格（即货币价值）. 如果可能，求平衡价格使每个部门的收支平衡.

表 1-1 一个简单的经济问题

部门的产出分配			采购部门
煤 炭	电 力	钢 铁	
0.0	0.4	0.6	煤炭
0.6	0.1	0.2	电力
0.4	0.5	0.2	钢铁

㊀ 见 Wassily W. Leontief, “Input-Output Economics,” *Scientific American*, October 1951, pp.15-21.



解 某一部门所在的一列表示它的产出的去向，它所在的一行表示它从哪些部门获得了投入. 例如，表 1-1 的第一行说明煤炭部门接受（采购）40%的电力产出和 60%的钢铁产出. 因为相应部门的总产出价格为 p_E和p_S，煤炭部门必须支付电力部门 $0.4p_E$ 美元，支付钢铁部门 $0.6p_S$ 美元，因此煤炭部门的总支出是 $0.4p_E+0.6p_S$ 美元. 为使煤炭部门的总收入 p_C 等于它的总支出,有

$$p_C = 0.4p_E + 0.6p_S \tag{1}$$

交易表的第二行说明电力部门的开支有 $0.6p_C$ 美元采购煤炭，$0.1p_E$ 美元采购电力，$0.2p_S$ 美元采购钢铁.因此电力部门的收支平衡条件是

$$p_E = 0.6p_C + 0.1p_E + 0.2p_S \tag{2}$$

最后，交易表的第三行导出最后的条件：

$$p_S = 0.4p_C + 0.5p_E + 0.2p_S \tag{3}$$

为求解方程组（1）、(2)、(3)，将所有未知量移到方程的左边并合并同类项. （例如，在方程（2）的左边将 $p_E - 0.1p_E$ 写成 $0.9p_E$.）

$$\begin{aligned} p_C - 0.4p_E - 0.6p_S &= 0 \\ -0.6p_C + 0.9p_E - 0.2p_S &= 0 \\ -0.4p_C - 0.5p_E + 0.8p_S &= 0 \end{aligned}$$

接下来进行行简化. 为简明起见，数值舍入到小数点后两位.

$$\begin{bmatrix} 1 & -0.4 & -0.6 & 0 \\ -0.6 & 0.9 & -0.2 & 0 \\ -0.4 & -0.5 & 0.8 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -0.4 & -0.6 & 0 \\ 0 & 0.66 & -0.56 & 0 \\ 0 & -0.66 & 0.56 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -0.4 & -0.6 & 0 \\ 0 & 0.66 & -0.56 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -0.4 & -0.60 & 0 \\ 0 & 1 & -0.85 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -0.94 & 0 \\ 0 & 1 & -0.85 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

通解是 $p_C = 0.94p_S, p_E = 0.85p_S$， p_S 为自由变量. 这个经济问题的平衡价格向量为

$$\boldsymbol{p} = \begin{bmatrix} p_C \\ p_E \\ p_S \end{bmatrix} = \begin{bmatrix} 0.94p_S \\ 0.85p_S \\ p_S \end{bmatrix} = p_S \begin{bmatrix} 0.94 \\ 0.85 \\ 1 \end{bmatrix}$$

任意（非负）p_S 取值可以算出平衡价格的一种取值. 例如，如果取 p_S 为 100（或 [illegible] 亿美元），那么 $p_C = 94, p_E = 85$.即如果煤炭的产出价格是 9 400 万美元，电力的产出价格是 8 500 万美元，钢铁的产出价格是 1 亿美元，那么每个部门的总收入和总支出将会相等. ■

配平化学方程式

化学方程式描述了化学反应的物质消耗和生产的数量. 例如，当丙烷气体燃烧时，丙烷（C_3H_8）与氧（O_2）结合生成二氧化碳（CO_2）和水（H_2O），按照如下形式的一个方程式

$$(x_1)C_3H_8 + (x_2)O_2 \to (x_3)CO_2 + (x_4)H_2O \tag{4}$$

为"配平"这个方程式，化学家必须找到 $x_1,\cdots,x_4$ 的全体数量，使得方程式左边碳（C）、氢（H）、氧（O）原子的总数等于右边相应原子的总数（因为在化学反应中原子既不会被破坏，也不会被创造）.

配平化学方程式的一个系统方法是建立描述一个化学反应中每一种类型的原子的数目的一个向量方程. 由于方程式（4）包含三种类型的原子（碳、氢、氧），给（4）式的每一种反应物和生成物构造一个属于 $\mathbb{R}^3$ 的向量，列出每个分子的组成原子的数目如下：

$$C_3H_8:\begin{bmatrix}3\\8\\0\end{bmatrix}, O_2:\begin{bmatrix}0\\0\\2\end{bmatrix}, CO_2:\begin{bmatrix}1\\0\\2\end{bmatrix}, H_2O:\begin{bmatrix}0\\2\\1\end{bmatrix}\begin{matrix}\leftarrow \text{碳}\\\leftarrow \text{氢}\\\leftarrow \text{氧}\end{matrix}$$

要配平方程式（4），$x_1,\cdots,x_4$ 的系数必须满足

$$x_1\begin{bmatrix}3\\8\\0\end{bmatrix}+x_2\begin{bmatrix}0\\0\\2\end{bmatrix}=x_3\begin{bmatrix}1\\0\\2\end{bmatrix}+x_4\begin{bmatrix}0\\2\\1\end{bmatrix}$$

将全部项移到等式左边（修改第三和第四个向量的符号），得到：

$$x_1\begin{bmatrix}3\\8\\0\end{bmatrix}+x_2\begin{bmatrix}0\\0\\2\end{bmatrix}+x_3\begin{bmatrix}-1\\0\\-2\end{bmatrix}+x_4\begin{bmatrix}0\\-2\\-1\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

行化简该方程组的增广矩阵得到通解

$$x_1 = \frac{1}{4}x_4, x_2 = \frac{5}{4}x_4, x_3 = \frac{3}{4}x_4, x_4\text{是自由变量}$$

因为化学方程式的系数应为整数，取 $x_4=4$, 那么$x_1 = 1, x_2 = 5, x_3 = 3$. 配平的方程式为

$$C_3H_8 + 5O_2 \to 3CO_2 + 4H_2O$$

如果方程式中的每个系数乘两倍（比如说），该方程式仍然是配平的. 然而在一般情形下，化学家倾向于使用全体系数尽可能小的数来配平方程式.

网络流

当科学家、工程师或经济学家研究一些数量在网络中的流动时自然推导出线性方程组. 例如，城市规划和交通工程人员监控一个网格状的市区道路的交通流量模式；电气工程师计算流经电路的电流；以及经济学家分析通过分销商和零售商的网络从制造商到顾客的产品销售. 许多网络中的方程组涉及成百甚至上千的变量和方程.

一个网络包含一组称为接合点或节点的点集，并由称为分支的线或弧连接部分或全部的节点. 流的方向在每个分支上有标示，流量（速度）也有显示或用变量标记.

网络流的基本假设是网络的总流入量等于总流出量，且流经一个节点的总输入等于总输出. 例如，图 1-28 显示 30 单元经过一个分支流入一个节点，x_1和x_2标记该节点经过其他分支的流出. 因为流量在每个节点中是守恒的，我们有 $x_1+x_2=30$. 类似地，每个节点的流量可以用一个方程描述. 网络分析的问题就是确定当局部信息（如网络的输入）已知时每一分支的流量.

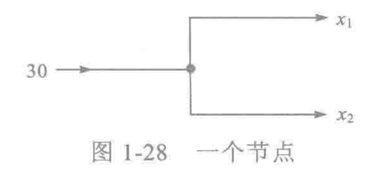


图 1-28　一个节点

例 2　图 1-29 中的网络是巴尔的摩市区一些单行道路在一个下午早些时候（以每小时车辆数目计算）的交通流量. 计算该网络的车流量.

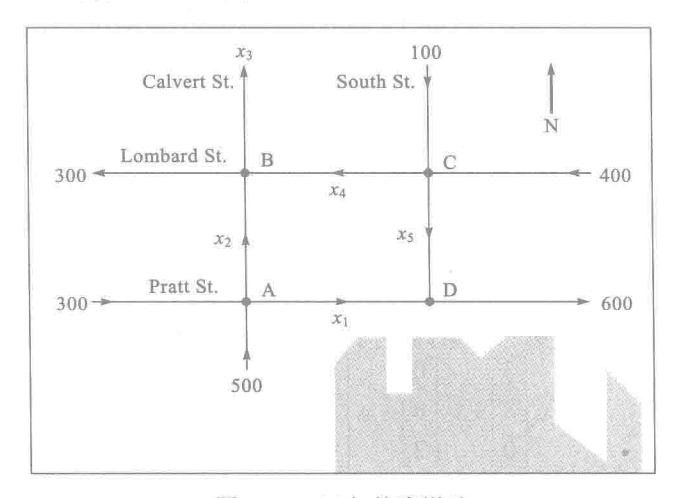


图 1-29　巴尔的摩道路

解　写出该流量的方程组，并求其通解. 如图 1-29 所示，标记道路交叉口（节点）和未知的分支流量. 在每个交叉口，令其车辆驶入数目等于车辆驶出数目.

交叉口	车辆驶入数目		车辆驶出数目
A	300+500	=	x_1+x_2
B	x_2+x_4	=	$300+x_3$
C	100+400	=	x_4+x_5
D	x_1+x_5	=	600

并且，网络中的总流入量（500+300+100+400）等于总流出量（300+ x_3 +600），经简化得 $x_3=400$. 该方程与上面四个方程联立并重排后得到下面的方程组：

$$x_1+x_2 \qquad\qquad =800$$

$$\begin{aligned} x_2 - x_3 + x_4 \quad &= 300 \\ x_4 + x_5 &= 500 \\ x_1 \quad + x_5 &= 600 \\ x_3 \quad &= 400 \end{aligned}$$

行化简相应的增广矩阵得到

$$\begin{aligned} x_1 \quad +x_5 &= 600 \\ x_2 \quad -x_5 &= 200 \\ x_3 \quad &= 400 \\ x_4 +x_5 &= 500 \end{aligned}$$

该网络的车流量为

$$\begin{cases} x_1 = 600 - x_5 \\ x_2 = 200 + x_5 \\ x_3 = 400 \\ x_4 = 500 - x_5 \\ x_5 \text{ 是自由变量} \end{cases}$$

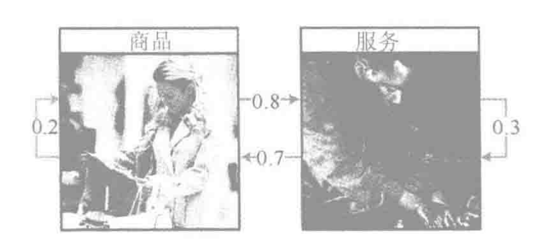
网络分支中的一个负流量对应模型中显示方向相反的流量. 由于本问题中的道路是单行线，这里不允许有负值变量. 这种情况给变量的可能取值增加了某种限制. 例如，因为 x_4 不能取负值，因此 $x_5 \leqslant 500$. 其他变量的约束在练习题 2 中有考虑. ■

练习题

1. 假设一个经济有农业、矿业和制造业三个部门. 农业部门销售它的产出的 5%给矿业部门，30%给制造业部门，保留余下的产出. 矿业部门销售它的产出的 20%给农业部门，70%给制造业部门，保留余下的产出. 制造业部门销售它的产出的 20%给农业部门，30%给矿业部门，保留余下的产出. 计算该经济的交易表，表中的列给出各个部门的产出如何分配给其他部门.
2. 考虑例 2 中的网络流. 确定 x_1 和 x_2 的可能取值范围. （提示：在例中 $x_5 \leqslant 500$. 这对 x_1 和 x_2 意味着什么？同时 $x_5 \geqslant 0$.）

习题 1.6

1. 假设一个经济只有商品和服务两个部门. 在每一年中，商品部门销售它的总产出的 80%给服务部门，并保留余下的产出，而服务部门销售它的总产出的 70%给商品部门，并保留余下的产出. 找出商品和服务部门的年度产出的平衡价格使得每一部门的收支平衡.



2. 找出例 1 中经济的另一组平衡价格.假设同样的经济中使用日元而不是美元来衡量各部门的产出值，讨论这个问题会有什么变化.
3. 考虑一个由燃料动力、制造和服务三个部门构成的经济体系. 燃料动力部门销售 80%的产出给制造部门和 10%的产出给服务部门，保留余下的产出. 制造部门销售 10%的产出给燃料动力部门和 80%的产出给服务部门，保留余下的产出. 服务部门销售 20%的产出给燃料动力部

门和 40%的产出给制造部门并保留余下的产出.

a. 写出该经济体系的交易表.

b. 建立方程组表示各部门收支平衡的条件.写出对应的增广矩阵以便行化简求平衡价格.

c. [M]找出当服务部门产出的价格是 100 个单位时的一组平衡价格.

4. 假设一个经济体系有四个部门，分别是矿业（M）、木材（L）、能源（E）和运输（T）部门. 部门 M 销售产出的 10%给部门 L 和 60%给部门 M 并保留余下的产出. 部门 L 销售产出的 15%给部门 M，50%给部门 E 和 20%给部门 T 并保留余下的产出. 部门 E 销售产出的 20%给部门 M，15%给部门 L 和 20%给部门 T 并保留余下的产出. 部门 T 销售产出的 20%给部门 M，10%给部门 L 和 50%给部门 E 并保留余下的产出.

a. 写出该经济体系的交易表.

b. [M]找出该经济体系的一组平衡价格.

5. 假设一个经济系统有四个行业构成，分别是农业、制造业、服务业和交通业. 农业产出的 20%销售给制造业、30%销售给服务业、30%销售给交通业，保留余下的产出. 制造业产出的 35%销售给农业、35%销售给服务业、20%销售给交通业，保留余下的产出. 服务业产出的 10%销售给农业、20%销售给制造业、20%销售给交通业，保留余下的产出. 交通业产出的 20%销售给农业、30%销售给制造业、20%销售给服务业，保留余下的产出.

a. 写出该经济系统的交易表；

b. [M]求出当交通业产出的单位价格是 10 美元时的平衡价格.

c. 服务业成功地推出一个“吃农场新鲜”活动，并增加其对农业的销售份额，增加到 40%，而农业产出将降低对制造业的销售份额，下降到 10%. 确定这个新经济系统的行情交易表.

d. [M]当交通业产出的单位价格是 10 美元时，求新的平衡价格. “吃农场新鲜”活动对该行业的平衡价格有什么影响?

习题 6~11 使用本节讨论的向量方程的方法配平化学方程式.

6. 氧化铝和碳发生反应生成铝和二氧化碳

$$Al_3O_8 + C \rightarrow Al + CO_2$$

（对每种化合物，构造一个向量列出铝、氧和碳的原子数目）.

7. Alka-Seltzer 碱性苏打包含重碳酸钠($NaHCO_3$)和柠檬酸($H_3C_6H_5O_7$). 当一颗药片溶解在水中时，会发生化学反应生成柠檬酸钠、水和二氧化碳（气体）:

$$NaHCO_3 + H_3C_6H_5O_7 \rightarrow Na_3C_6H_5O_7 + H_2O + CO_2$$

8. 碳酸钙（$CaCO_3$）中和酸（H_3O），发生反应产生酸雨，其未配平的化学反应式为

$$H_3O + CaCO_3 \rightarrow H_2O + Ca + CO_2$$

9. 硫化硼与水剧烈反应生成硼酸和硫化氢气体（臭蛋味）. 未配平的化学反应式为

$$B_2S_3 + H_2O \rightarrow H_3BO_3 + H_2S$$

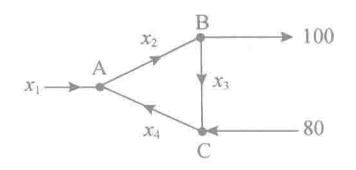
10. [M]如果可能，使用精确的算术或合理的计算格式配平如下的化学反应方程式:

$$PbN_6 + CrMn_2O_8 \rightarrow Pb_3O_4 + Cr_2O_3 + MnO_2 + NO$$

11. [M]下面的化学反应可以在工业过程中应用，如砷（AsH_3）的生产. 配平方程式.

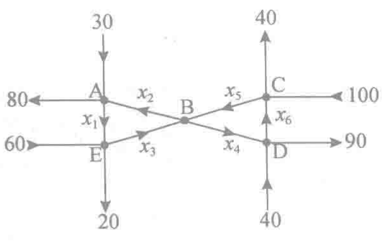
$$MnS + As_2Cr_{10}O_{35} + H_2SO_4 \rightarrow HMnO_4 + AsH_3 + CrS_3O_{12} + H_2O$$

12. 求下图中网络流量的通解. 假设流量都是非负的，x_4 可能的最小值是什么?



13. a. 求下图中网络的交通流量的通解.

b. 假设流量必须以标示的方向流动，求分支 x_2, x_3, x_4, x_5 的流量的最小值.

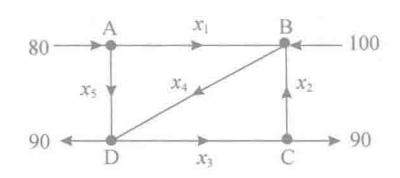


14. a. 求下图中高速公路网络的交通流量的通解.

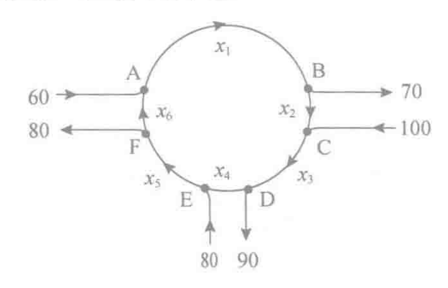
（流量以车辆数/分钟计算.）

b. 求 x_5 的道路交通封闭时的交通流量的通解.

c. 当 $x_5=0$ 时，x_4 的最小值是什么？



15. 英格兰的交叉路通常被设计成单行的“环行路”，如下图所示. 假设流量必须以标示的方向流动，求 x_6 的最小值.



练习题答案

1. 将比例用小数表示，由于要考虑所有的产出，每列的元素之和等于 1. 这种情况有助于填补空缺的元素.

部门的产出分配			采购部门
农业	矿业	制造业	
0.65	0.20	0.20	农业
0.05	0.10	0.30	矿业
0.30	0.70	0.50	制造业

2. 因 $x_1\leqslant 500$，由 x_1 和 x_2 的方程得到 $x_1\geqslant 100$ 和 $x_2\leqslant 700$，由 $x_5\geqslant 0$ 得到 $x_1\leqslant 600$ 和 $x_2\geqslant 200$. 因此 $100\leqslant x_1\leqslant 600$ 和 $200\leqslant x_2\leqslant 700$.

1.7 线性无关

1.5 节的齐次线性方程组也可从另一观点研究，即把它们写成向量方程. 这时，重点从 $A\boldsymbol{x}=\boldsymbol{0}$ 的未知数的解转向出现在向量方程中的向量.

例如，考虑方程

$$x_1\begin{bmatrix}1\\2\\3\end{bmatrix}+x_2\begin{bmatrix}4\\5\\6\end{bmatrix}+x_3\begin{bmatrix}2\\1\\0\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix} \tag{1}$$

此方程当然有平凡解，即 $x_1=x_2=x_3=0$，如 1.5 节，主要问题是，平凡解是否是唯一解.

定义 $\mathbb{R}^n$ 中一组向量 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 称为**线性无关**的，若向量方程

$$x_1\boldsymbol{v}_1+x_2\boldsymbol{v}_2+\cdots+x_p\boldsymbol{v}_p=\boldsymbol{0}$$

仅有平凡解. 向量组（集）$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 称为**线性相关**的，若存在不全为零的权 $c_1,\cdots,c_p$，使

$$c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_p\boldsymbol{v}_p=\boldsymbol{0} \tag{2}$$

方程（2）称为向量 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 之间的**线性相关关系**，其中权不全为零. 一组向量为线性相关，当且仅当它不是线性无关的. 为简单起见，我们也可说 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 线性相关，意思是向量组（集）

$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$是线性相关组. 对线性无关组也使用类似的语言.

例 1 设

$$\boldsymbol{v}_1=\begin{bmatrix}1\\2\\3\end{bmatrix},\ \boldsymbol{v}_2=\begin{bmatrix}4\\5\\6\end{bmatrix},\ \boldsymbol{v}_3=\begin{bmatrix}2\\1\\0\end{bmatrix}$$

a. 确定向量组$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$是否线性相关.

b. 可能的话，求出$\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$的一个线性相关关系.

解 a. 我们需要确定方程（1）是否有非平凡解，把增广矩阵进行行变换，得

$$\begin{bmatrix}1&4&2&0\\2&5&1&0\\3&6&0&0\end{bmatrix}\sim\begin{bmatrix}1&4&2&0\\0&-3&-3&0\\0&0&0&0\end{bmatrix}$$

显然，x_1和x_2为基本变量，x_3为自由变量. x_3的每个非零值确定（1）的一组非平凡解，因此$\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$线性相关.

b. 为求出$\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$的线性关系，继续化简增广矩阵，写出新的方程组

$$\begin{bmatrix}1&0&-2&0\\0&1&1&0\\0&0&0&0\end{bmatrix}\qquad\begin{aligned}x_1\qquad-2x_3&=0\\x_2+\ x_3&=0\\0&=0\end{aligned}$$

故$x_1=2x_3,x_2=-x_3,x_3$为自由变量. 选取x_3的一个非零值，比如$x_3=5$，则$x_1=10,x_2=-5$，把这些值代入（2）得

$$10\boldsymbol{v}_1-5\boldsymbol{v}_2+5\boldsymbol{v}_3=\boldsymbol{0}$$

这是一个（无穷多个之中的一个）可能的线性相关关系. ■

矩阵各列的线性无关

设我们考虑矩阵$A=[\boldsymbol{a}_1\cdots\boldsymbol{a}_n]$，矩阵方程$A\boldsymbol{x}=\boldsymbol{0}$可以写成

$$x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+\cdots+x_n\boldsymbol{a}_n=\boldsymbol{0}$$

A的各列之间的每一个线性相关关系对应于方程$A\boldsymbol{x}=\boldsymbol{0}$的一个非平凡解. 因此我们有下列重要事实.

> 矩阵A的各列线性无关，当且仅当方程$A\boldsymbol{x}=\boldsymbol{0}$仅有平凡解.

例 2 确定矩阵$A=\begin{bmatrix}0&1&4\\1&2&-1\\5&8&0\end{bmatrix}$的各列是否线性无关.

解 研究$A\boldsymbol{x}=\boldsymbol{0}$，把增广矩阵行变换：

$$\begin{bmatrix}0&1&4&0\\1&2&-1&0\\5&8&0&0\end{bmatrix}\sim\begin{bmatrix}1&2&-1&0\\0&1&4&0\\0&-2&5&0\end{bmatrix}\sim\begin{bmatrix}1&2&-1&0\\0&1&4&0\\0&0&13&0\end{bmatrix}$$

此时，显然方程有 3 个基本变量，没有自由变量，因此方程 $A\boldsymbol{x}=\boldsymbol{0}$ 仅有平凡解，A 的各列是线性无关的. ■

一个或两个向量的集合

仅含一个向量，比如说由 $\boldsymbol{v}$ 形成的集合线性无关当且仅当 $\boldsymbol{v}$ 不是零向量. 这是因为当 $\boldsymbol{v}\neq\boldsymbol{0}$ 时向量方程 $x_1\boldsymbol{v}=\boldsymbol{0}$ 仅有平凡解. 零向量是线性相关的，因 $x_1\boldsymbol{0}=\boldsymbol{0}$ 有许多非平凡解.

下列例子说明两个向量线性相关的情况.

例 3 确定下列向量组是否线性无关

a. $\boldsymbol{v}_1=\begin{bmatrix}3\\1\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}6\\2\end{bmatrix}$ b. $\boldsymbol{v}_1=\begin{bmatrix}3\\2\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}6\\2\end{bmatrix}$

解 a. 注意 $\boldsymbol{v}_2$ 是 $\boldsymbol{v}_1$ 的倍数，即 $\boldsymbol{v}_2=2\boldsymbol{v}_1$. 因此$-2\boldsymbol{v}_1+\boldsymbol{v}_2=\boldsymbol{0}$，表明 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 线性相关.

b. $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的任意一个不是另一个的倍数. 它们能否线性相关？设 c 和 d 满足

$$c\boldsymbol{v}_1+d\boldsymbol{v}_2=\boldsymbol{0}$$

若 $c\neq 0$，我们可用 $\boldsymbol{v}_2$ 表示 $\boldsymbol{v}_1$，即 $\boldsymbol{v}_1=(-d/c)\boldsymbol{v}_2$，这是不可能的，因 $\boldsymbol{v}_1$ 不是 $\boldsymbol{v}_2$ 的倍数. 故 c 必是零. 类似地 d 必是 0，于是 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是线性无关组. ■

例 3 中的讨论说明，你总可以用观察法来决定两个向量是否线性相关. 行变换是不必要的，只要看一个向量是否是另一个向量的倍数即可（这个方法只能用于两个向量的情况）.

> 两个向量的集合 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 线性相关，当且仅当其中一个向量是另一个向量的倍数. 这个集合线性无关，当且仅当其中任一个向量都不是另一个向量的倍数.

从几何意义上看，两个向量线性相关，当且仅当它们落在通过原点的同一条直线上. 图 1-30 表示例 3 中两组向量的情况.

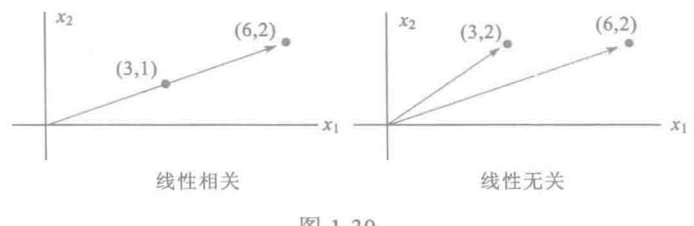


图 1-30

两个或更多个向量的集合

下面定理的证明类似于例 3 的思路. 详细的证明在本节末给出.

定理 7 （线性相关集的特征）

两个或更多个向量的集合 $S=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 线性相关，当且仅当 S 中至少有一个向量是其他向量的线性组合，事实上，若 S 线性相关，且 $\boldsymbol{v}_1\neq 0$，则某个 $\boldsymbol{v}_j(j>1)$ 是它前面几个向量 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_{j-1}$ 的线性组合.

警告 定理 7 没有说在线性相关集中每一个向量都是它前面的向量的线性组合，线性

相关集中某个向量可能不是其他向量的线性组合. 见练习题 3.

例 4 设 $\boldsymbol{u}=\begin{bmatrix}3\\1\\0\end{bmatrix},\boldsymbol{v}=\begin{bmatrix}1\\6\\0\end{bmatrix}$，叙述 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 生成的集合，并说明向量 $\boldsymbol{w}$ 属于 $\text{Span}\{\boldsymbol{u},\boldsymbol{v}\}$ 当且仅当 $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$ 线性相关.

解 向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是线性无关的，因为它们之中任何一个不是另一个的倍数，所以它们生成 $\mathbb{R}^3$ 中一个平面（见 1.3 节），事实上，$\text{Span}\{\boldsymbol{u},\boldsymbol{v}\}$ 就是 x_1x_2 平面（即 $x_3=0$），若 $\boldsymbol{w}$ 是 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的线性组合，由定理 7 知 $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$ 线性相关，反之，设 $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$ 线性相关，由定理 7 知 $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$ 中某一向量是它前面的向量的线性组合（因 $\boldsymbol{u}\neq\boldsymbol{0}$），这向量必是 $\boldsymbol{w}$，因为 $\boldsymbol{v}$ 不是 $\boldsymbol{u}$ 的倍数. 因而 $\boldsymbol{w}$ 属于 $\text{Span}\{\boldsymbol{u},\boldsymbol{v}\}$，见图 1-31.

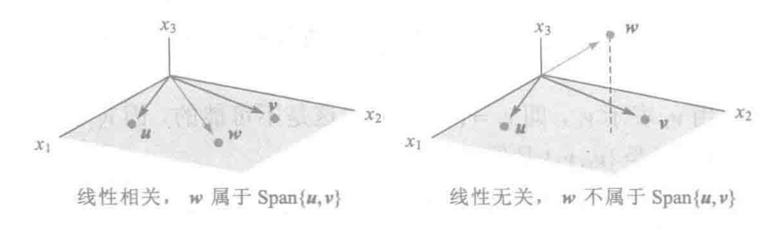


图 1-31 $\mathbb{R}^3$ 中的线性相关性 ■

例 4 可推广到 $\mathbb{R}^3$ 中任意集合 $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$，其中 $\boldsymbol{u}$ 与 $\boldsymbol{v}$ 线性无关. 这时集合 $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$ 线性相关当且仅当 $\boldsymbol{w}$ 在 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 所生成的平面上.

下面两个定理说明了线性相关的一些条件. 定理 8 在今后各章中是一个关键的结果.

定理 8 若一个向量组的向量个数超过每个向量元素个数，那么这个向量组线性相关. 就是说，$\mathbb{R}^n$ 中任意向量组 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$，当 $p>n$ 时线性相关.

证 设 $A=[\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p]$，则 A 是 $n\times p$ 矩阵，方程对应于 p 个未知量的 n 个方程，若 $p>n$，未知量比方程多，所以必定有自由变量. 因此 $A\boldsymbol{x}=\boldsymbol{0}$ 必有非平凡解，所以 A 的各列线性相关，图 1-32 给出了这个定理的矩阵说明.

$$n\overset{p}{\begin{bmatrix}* & * & * & * & *\\ * & * & * & * & *\\ * & * & * & * & *\end{bmatrix}}$$

图 1-32 若 $p>n$，矩阵各列线性相关 ■

警告 定理 8 没有涉及向量组中向量个数不超过每个向量中元素个数的情形.

例 5 向量 $\begin{bmatrix}2\\1\end{bmatrix},\begin{bmatrix}4\\-1\end{bmatrix},\begin{bmatrix}-2\\2\end{bmatrix}$ 线性相关. 因为每个向量仅有 2 个元素而共有 3 个向量，注意：

其中任何一个向量并不是另一向量的倍数. 见图 1-33.

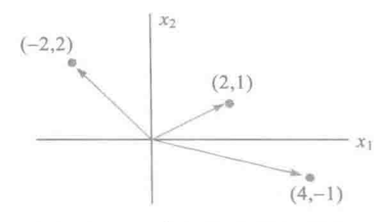


图 1-33 $\mathbb{R}^2$ 中的线性相关集 ■

定理 9 若向量组 $S=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 包含零向量，则它线性相关.

证 把这些向量重新编号，我们可设 $\boldsymbol{v}_1=0$，于是方程 $1\cdot\boldsymbol{v}_1+0\cdot\boldsymbol{v}_2+\cdots+0\cdot\boldsymbol{v}_p=\boldsymbol{0}$ 证明了 S 线性相关. ■

例 6 用观察法确定下列向量组是否线性相关.

a. $\begin{bmatrix}1\\7\\6\end{bmatrix},\begin{bmatrix}2\\0\\9\end{bmatrix},\begin{bmatrix}3\\1\\5\end{bmatrix},\begin{bmatrix}4\\1\\8\end{bmatrix}$ b. $\begin{bmatrix}2\\3\\5\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}1\\1\\8\end{bmatrix}$ c. $\begin{bmatrix}-2\\4\\6\\10\end{bmatrix},\begin{bmatrix}3\\-6\\-9\\15\end{bmatrix}$

解 a. 这个向量组包含 4 个向量，每个向量仅有 3 个元素，因此它们线性相关.

b. 定理 8 不能应用，因为向量个数不超过每个向量中元素个数. 因该组中有零向量，根据定理 9，因此它线性相关.

c. 若我们比较两向量的对应元素，第 2 个向量看来是第一个向量的−3/2 倍. 这个关系对前 3 对元素成立，但对第 4 个不成立. 因此，这两向量中任意一个不是另一个的倍数，因此是线性无关的. ■

一般地，你必须把每一节完整地读几遍才能理解像线性相关这样的重要概念. 学习指导书中的这一节的注解对你掌握线性代数中的这一重要思想是很重要的. 例如，下列的证明值得一读，因为它指出如何应用线性无关的定义.

定理 7（线性相关集的特征）的证明 若 S 中某个 $\boldsymbol{v}_j$ 是其他向量的线性组合，那么把方程两边减去 $\boldsymbol{v}_j$ 就产生一个线性关系，其中 $\boldsymbol{v}_j$ 的权为（−1），例如，若 $\boldsymbol{v}_1=c_2\boldsymbol{v}_2+c_3\boldsymbol{v}_3$，那么

$$\boldsymbol{0}=(-1)\boldsymbol{v}_1+c_2\boldsymbol{v}_2+c_3\boldsymbol{v}_3+0\boldsymbol{v}_4+\cdots+0\boldsymbol{v}_p$$

于是 S 线性相关.

反之，设 S 线性相关，若 $\boldsymbol{v}_1$ 为零，则它是 S 中其他向量的一个（平凡）线性组合. 若不为零，存在 $c_1,\cdots,c_p$ 不全为零，使

$$c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_p\boldsymbol{v}_p=\boldsymbol{0}$$

设 j 是使 $c_j\neq0$ 的最大下标. 若 $j=1$ 则 $c_1\boldsymbol{v}_1=\boldsymbol{0}$，这是不可能的，因 $\boldsymbol{v}_1\neq0$，故 $j>1$. 而

$$c_1\boldsymbol{v}_1+\cdots+c_j\boldsymbol{v}_j+0\boldsymbol{v}_{j+1}+\cdots+0\boldsymbol{v}_p=\boldsymbol{0}$$

$$c_j\boldsymbol{v}_j=-c_1\boldsymbol{v}_1-\cdots-c_{j-1}\boldsymbol{v}_{j-1}$$

$$\boldsymbol{v}_j=\left(-\frac{c_1}{c_j}\right)\boldsymbol{v}_1+\cdots+\left(-\frac{c_{j-1}}{c_j}\right)\boldsymbol{v}_{j-1}$$

■

练习题

设 $\boldsymbol{u}=\begin{bmatrix}3\\2\\-4\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}-6\\1\\7\end{bmatrix}$，$\boldsymbol{w}=\begin{bmatrix}0\\-5\\2\end{bmatrix}$，$\boldsymbol{z}=\begin{bmatrix}3\\7\\-5\end{bmatrix}$.

1. 集合 $\{\boldsymbol{u},\boldsymbol{v}\},\{\boldsymbol{u},\boldsymbol{w}\},\{\boldsymbol{u},\boldsymbol{z}\},\{\boldsymbol{v},\boldsymbol{w}\},\{\boldsymbol{v},\boldsymbol{z}\}$ 和 $\{\boldsymbol{w},\boldsymbol{z}\}$ 都是线性无关的吗？为什么？
2. 练习题 1 的答案是否蕴涵着 $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w},\boldsymbol{z}\}$ 也线性无关？
3. 为确定 $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w},\boldsymbol{z}\}$ 是否线性相关，是否有必要验证 $\boldsymbol{w}$ 是 $\boldsymbol{u},\boldsymbol{v},\boldsymbol{z}$ 的线性组合？
4. $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w},\boldsymbol{z}\}$ 是否线性相关？

习题 1.7

习题 1~4 中，确定向量组是否线性相关，给出理由.

1. $\begin{bmatrix}5\\0\\0\end{bmatrix},\begin{bmatrix}7\\2\\-6\end{bmatrix},\begin{bmatrix}9\\4\\-8\end{bmatrix}$

2. $\begin{bmatrix}0\\2\\3\end{bmatrix},\begin{bmatrix}0\\5\\-8\end{bmatrix},\begin{bmatrix}-1\\3\\1\end{bmatrix}$

3. $\begin{bmatrix}2\\-3\end{bmatrix},\begin{bmatrix}-4\\6\end{bmatrix}$

4. $\begin{bmatrix}-1\\3\end{bmatrix},\begin{bmatrix}-3\\-9\end{bmatrix}$

习题 5~8 中，确定给定矩阵的各列是否构成线性相关集，给出理由.

5. $\begin{bmatrix}0&-3&9\\2&1&-7\\-1&4&-5\\1&-4&-2\end{bmatrix}$

6. $\begin{bmatrix}-4&-3&0\\0&-1&5\\1&1&-5\\2&1&-10\end{bmatrix}$

7. $\begin{bmatrix}1&4&-3&0\\-2&-7&5&1\\-4&-5&7&5\end{bmatrix}$

8. $\begin{bmatrix}1&-2&3&2\\-2&4&-6&2\\0&1&-1&3\end{bmatrix}$

习题 9~10 中，（a）对 h 的什么值，$\boldsymbol{v}_3$ 属于 $\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$，（b）对 h 的什么值，$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 线性相关？给出理由.

9. $\boldsymbol{v}_1=\begin{bmatrix}1\\-3\\2\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}-3\\9\\-6\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}5\\-7\\h\end{bmatrix}$

10. $\boldsymbol{v}_1=\begin{bmatrix}1\\-3\\-5\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}-3\\9\\15\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}2\\-5\\h\end{bmatrix}$

习题 11~14 中，求出 h 的值，使向量组线性相关. 给出理由.

11. $\begin{bmatrix}2\\-2\\4\end{bmatrix},\begin{bmatrix}4\\-6\\7\end{bmatrix},\begin{bmatrix}-2\\2\\h\end{bmatrix}$

12. $\begin{bmatrix}3\\-6\\1\end{bmatrix},\begin{bmatrix}-6\\4\\-3\end{bmatrix},\begin{bmatrix}9\\h\\3\end{bmatrix}$

13. $\begin{bmatrix}1\\5\\-3\end{bmatrix},\begin{bmatrix}-2\\-9\\6\end{bmatrix},\begin{bmatrix}3\\h\\-9\end{bmatrix}$

14. $\begin{bmatrix}1\\-2\\-4\end{bmatrix},\begin{bmatrix}-3\\7\\6\end{bmatrix},\begin{bmatrix}2\\1\\h\end{bmatrix}$

习题 15~20 中，通过观察判断向量组是否线性无关，给出理由.

15. $\begin{bmatrix}5\\1\end{bmatrix},\begin{bmatrix}2\\8\end{bmatrix},\begin{bmatrix}1\\3\end{bmatrix},\begin{bmatrix}-1\\7\end{bmatrix}$

16. $\begin{bmatrix}2\\-4\\8\end{bmatrix},\begin{bmatrix}-3\\6\\-12\end{bmatrix}$

17. $\begin{bmatrix}5\\-3\\-1\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}-7\\2\\4\end{bmatrix}$

18. $\begin{bmatrix}3\\4\end{bmatrix},\begin{bmatrix}-1\\5\end{bmatrix},\begin{bmatrix}3\\5\end{bmatrix},\begin{bmatrix}7\\1\end{bmatrix}$

19. $\begin{bmatrix}-8\\12\\-4\end{bmatrix},\begin{bmatrix}2\\-3\\-1\end{bmatrix}$

20. $\begin{bmatrix}1\\4\\-7\end{bmatrix},\begin{bmatrix}-2\\5\\3\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix}$

习题 21~22 中，判断每个命题的真假. 仔细读课文，说明你的理由.

21. a. 若方程 $A\boldsymbol{x}=\boldsymbol{0}$ 有平凡解，则矩阵 A 的各列线性无关.

b. 若 S 是线性相关集，则 S 中每个向量是其他向量的线性组合.

c. 任意 4×5 矩阵的各列线性相关.

d. 若 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 线性无关，而 $\{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\}$ 线性相关，则 $\boldsymbol{z}$ 属于 $\text{Span}\{\boldsymbol{x},\boldsymbol{y}\}$.

22. a. 若 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 线性无关，而 $\boldsymbol{w}$ 属于 $\text{Span}\{\boldsymbol{u},\boldsymbol{v}\}$，则 $\{\boldsymbol{x},\boldsymbol{v},\boldsymbol{w}\}$ 线性相关.

b. 若 $\mathbb{R}^3$ 中三个向量位于 $\mathbb{R}^3$ 中同一个平面上，则这三个向量线性相关.

c. 某一个向量集的向量个数少于每个向量所含元素个数，则它线性无关.

d. 若 $\mathbb{R}^n$ 中一个向量集线性相关，则此集包含的向量个数大于 n.

习题 23~26 中，给出矩阵可能的阶梯形. 使用 1.2 节例 1 中的记号.

23. A 是 2×2 矩阵，两列线性相关.

24. A 是 3×3 矩阵，各列线性无关.

25. A 是 4×2 矩阵，$A=[\boldsymbol{a}_1\quad \boldsymbol{a}_2]$，$\boldsymbol{a}_2$ 不是 $\boldsymbol{a}_1$ 的倍数.

26. A 是 4×3 矩阵，$A=[\boldsymbol{a}_1\quad \boldsymbol{a}_2\quad \boldsymbol{a}_3]$，$\{\boldsymbol{a}_1,\boldsymbol{a}_2\}$ 线性无关，$\boldsymbol{a}_3$ 不属于 $\text{Span}\{\boldsymbol{a}_1,\boldsymbol{a}_2\}$.

27. 一 6×4 矩阵如果各列线性无关，则它有多少个主元列？为什么？

28. 一 4×6 矩阵如果其列生成 $\mathbb{R}^4$，则它有多少个主元列？为什么？

29. 构造 3×2 矩阵 A 和 B 使 $A\boldsymbol{x}=\boldsymbol{0}$ 仅有平凡解，$B\boldsymbol{x}=\boldsymbol{0}$ 有非平凡解.

30. a. 填空：“若 A 是 $m\times n$ 矩阵，则 A 的各列为线性无关，当且仅当 A 有________个主元列.”

b. 说明命题（a）为什么是真的.

习题 31~32 中，不使用行变换求解.（提示：将 $A\boldsymbol{x}=\boldsymbol{0}$ 写成向量方程.）

31. 给定 $A=\begin{bmatrix}2&3&5\\-5&1&-4\\-3&-1&-4\\1&0&1\end{bmatrix}$，观察到第 3 列是前两列之和. 求出 $A\boldsymbol{x}=\boldsymbol{0}$ 的一个非平凡解.

32. 给定 $A=\begin{bmatrix}4&3&-5\\-2&-2&4\\-2&-3&7\end{bmatrix}$，观察到第 1 列加上第 2 列的两倍等于第 3 列. 求 $A\boldsymbol{x}=\boldsymbol{0}$ 的一个非平凡解.

习题 33~38 的每个命题是（在任何情形下）真或（至少一种情形下）假. 若是假，举例说明该命题不是永远为真. 这样一个例子称为该命题的一个反例. 若是真，给出理由（一个特例不能说明该命题总是真的）. 这里你必须做比习题 21~22 更多的工作.

33. 若 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_4$ 属于 $\mathbb{R}^4$，$\boldsymbol{v}_3=2\boldsymbol{v}_1+\boldsymbol{v}_2$，则 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4\}$ 线性相关.

34. 若 $\boldsymbol{v}_1$ 与 $\boldsymbol{v}_2$ 属于 $\mathbb{R}^4$，$\boldsymbol{v}_2$ 不是 $\boldsymbol{v}_1$ 的倍数，则 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 线性无关.

35. 若 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_5$ 属于 $\mathbb{R}^5$，且 $\boldsymbol{v}_3=\boldsymbol{0}$，则 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4,\boldsymbol{v}_5\}$ 线性相关.

36. 若 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 属于 $\mathbb{R}^3$，而 $\boldsymbol{v}_3$ 不是 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 的线性组合，则 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 线性无关.

37. 若 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_4$ 属于 $\mathbb{R}^4$，$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 线性相关，则 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4\}$ 也线性相关.

38. 若 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_4$ 是 $\mathbb{R}^4$ 中线性无关向量，则 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 也线性无关.（提示：考虑方程 $x_1\boldsymbol{v}_1+x_2\boldsymbol{v}_2+x_3\boldsymbol{v}_3+0\cdot\boldsymbol{v}_4=\boldsymbol{0}$.）

39. 设 A 是 $m\times n$ 矩阵，且对 $\mathbb{R}^m$ 中所有 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 至多只有一个解. 应用线性无关的定义说明 A 的各列必定线性无关.

40. 设 $m\times n$ 矩阵 A 有 n 个主元列. 说明为什么对 $\mathbb{R}^m$ 中每个 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 至多有一个解.（提示：说明为什么 $A\boldsymbol{x}=\boldsymbol{b}$ 不能有无穷多个解）

[M]习题 41~42 中，求出 A 中尽可能多的列构造矩阵 B，使方程 $B\boldsymbol{x}=\boldsymbol{0}$ 仅有平凡解. 解方程 $B\boldsymbol{x}=\boldsymbol{0}$ 来证明你的结论.

41. $A=\begin{bmatrix}3&-4&10&7&-4\\-5&-3&-7&-11&15\\4&3&5&2&1\\8&-7&23&4&15\end{bmatrix}$

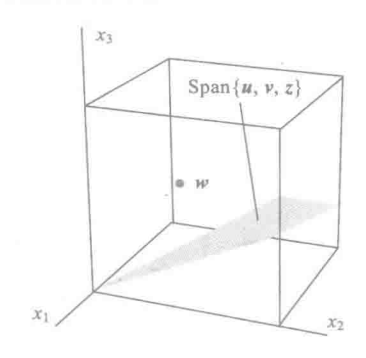
42. $A=\begin{bmatrix}12 & 10 & -6 & 8 & 4 & -14\\ -7 & -6 & 4 & -5 & -7 & 9\\ 9 & 9 & -9 & 9 & 9 & -18\\ -4 & -3 & -1 & 0 & -8 & 1\\ 8 & 7 & -5 & 6 & 1 & -11\end{bmatrix}$

43. [M]对习题 41 中的矩阵 A 和 B，选择 A 中未在构造 B 中使用的列 $\boldsymbol{v}$，确定 $\boldsymbol{v}$ 是否属于 B 的各列所生成的集（说明你的计算）.

44. [M]对习题 42 中的矩阵 A 和 B 重复习题 43 中的工作，说明你发现了什么.

练习题答案

1. 是，每一种情况下，任一个向量都不是另一个的倍数. 所以每个向量组都线性无关.
2. 否，练习题 1 中的观察并不说明 $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w},\boldsymbol{z}\}$ 线性无关.
3. 否，检验线性无关性，去检验某个特定向量是否其他向量的线性组合不是好方法. 有可能某一向量不是其他向量的线性组合而整个向量组仍然线性相关，本题中，$\boldsymbol{w}$ 不是 $\boldsymbol{u},\boldsymbol{v},\boldsymbol{z}$ 的线性组合.
4. 是，由定理 8，向量个数（4）超过元素的个数（3）.



1.8 线性变换介绍

矩阵方程 $A\boldsymbol{x}=\boldsymbol{b}$ 和对应的向量方程 $x_1\boldsymbol{a}_1+\cdots+x_n\boldsymbol{a}_n=\boldsymbol{b}$ 之间的差别仅仅是记号上的不同，然而，矩阵方程还可以由线性代数中其他问题以及应用（例如计算机图形、信号传递等）所引起，而不仅是直接与向量线性组合的问题有关. 通常的情况是把矩阵 A 当作一种“对象”，它通过乘法“作用”于向量 $\boldsymbol{x}$，产生的新向量称为 $A\boldsymbol{x}$.

例如，方程

$$\underset{A}{\begin{bmatrix}4 & -3 & 1 & 3\\ 2 & 0 & 5 & 1\end{bmatrix}}\underset{\boldsymbol{x}}{\begin{bmatrix}1\\1\\1\\1\end{bmatrix}}=\underset{\boldsymbol{b}}{\begin{bmatrix}5\\8\end{bmatrix}} \text{和} \underset{A}{\begin{bmatrix}4 & -3 & 1 & 3\\ 2 & 0 & 5 & 1\end{bmatrix}}\underset{\boldsymbol{u}}{\begin{bmatrix}1\\4\\-1\\3\end{bmatrix}}=\underset{\boldsymbol{0}}{\begin{bmatrix}0\\0\end{bmatrix}}$$

乘以矩阵 A 后，将 $\boldsymbol{x}$ 变成 $\boldsymbol{b}$，将 $\boldsymbol{u}$ 变成零向量，见图 1-34.

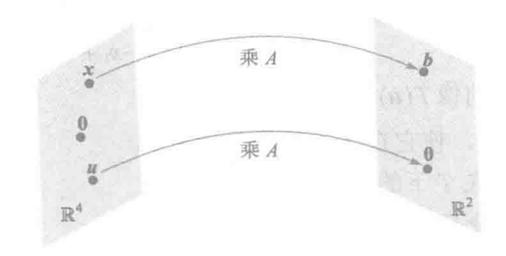


图 1-34　通过矩阵乘法变换向量

由这个新观点，解方程 $A\boldsymbol{x}=\boldsymbol{b}$ 就是要求出 $\mathbb{R}^4$ 中所有经过乘以 A 的“作用”后变为 $\boldsymbol{b}$ 的向量 $\boldsymbol{x}$.

由 $\boldsymbol{x}$ 到 $A\boldsymbol{x}$ 的对应是由一个向量集到另一个向量集的函数. 这个概念推广了通常的函数概念，通常的函数是把一个实数变为另一个实数的规则.

由 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的一个**变换**（或称**函数**、**映射**）T 是一个规则，它把 $\mathbb{R}^n$ 中每个向量 x 对应以 $\mathbb{R}^m$ 中的一个向量 $T(\boldsymbol{x})$. 集 $\mathbb{R}^n$ 称为 T 的**定义域**，而 $\mathbb{R}^m$ 称为 T 的**余定义域**（或取值空间）. 符号 T : $\mathbb{R}^n\to\mathbb{R}^m$ 说明 T 的定义域是 $\mathbb{R}^n$ 而余定义域是 $\mathbb{R}^m$，对于 $\mathbb{R}^n$ 中向量 $\boldsymbol{x}$，$\mathbb{R}^m$ 中向量 $T(\boldsymbol{x})$ 称为 $\boldsymbol{x}$（在 T 作用下）的**像**. 所有像 $T(\boldsymbol{x})$ 的集合称为 T 的**值域**. 见图 1-35.

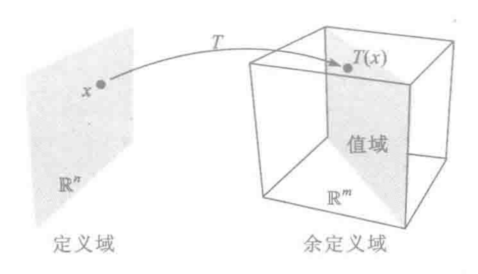


图 1-35　$T:\mathbb{R}^n\to\mathbb{R}^m$ 的定义域、余定义域和值域

本节的新名词是非常重要的，因为矩阵乘法的动态观点对理解线性代数和建立时变物理系统的数学模型起着关键作用. 这样的*动力系统*将在 1.10 节、4.8 节、4.9 节以及第 5 章进行讨论.

矩阵变换

本节以下内容研究有关矩阵乘法的映射. 对 $\mathbb{R}^n$ 中每个 $\boldsymbol{x}$，$T(\boldsymbol{x})$ 由 $A\boldsymbol{x}$ 计算得到，其中 A 是 $m\times n$ 矩阵，为简单起见，有时将这样一个*矩阵变换*记为 $\boldsymbol{x}\to A\boldsymbol{x}$，注意当 A 有 n 列时，T 的定义域为 $\mathbb{R}^n$，而当 A 的每个列有 m 个元素时，T 的余定义域为 $\mathbb{R}^m$. T 的值域为 A 的列的所有线性组合的集合，因为每个像 $T(\boldsymbol{x})$ 有 $A\boldsymbol{x}$ 的形式.

例 1　设 $A=\begin{bmatrix}1&-3\\3&5\\-1&7\end{bmatrix},\boldsymbol{u}=\begin{bmatrix}2\\-1\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}3\\2\\-5\end{bmatrix},\boldsymbol{c}=\begin{bmatrix}3\\2\\5\end{bmatrix}$，定义变换 $T:\mathbb{R}^2\to\mathbb{R}^3$ 为 $T(\boldsymbol{x})=A\boldsymbol{x}$，于是

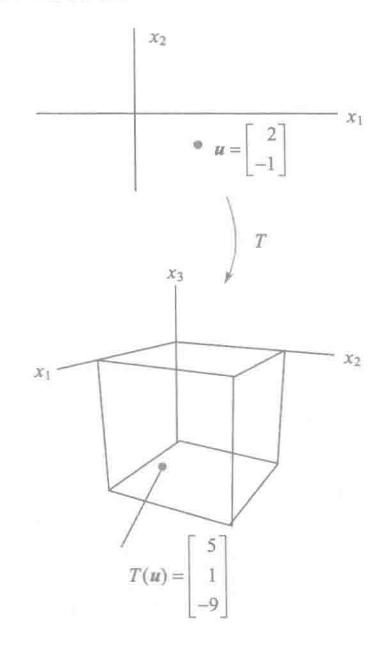
$$T(\boldsymbol{x})=A\boldsymbol{x}=\begin{bmatrix}1&-3\\3&5\\-1&7\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}x_1-3x_2\\3x_1+5x_2\\-x_1+7x_2\end{bmatrix}$$

a. 求 $\boldsymbol{u}$ 在变换 T 下的像 $T(\boldsymbol{u})$.

b. 求 $\mathbb{R}^2$ 中的向量 $\boldsymbol{x}$，使它在 T 下的像是向量 $\boldsymbol{b}$.

c. 是否有其他向量在 T 下的像也是 $\boldsymbol{b}$？

d. 确定 $\boldsymbol{c}$ 是否属于变换 T 的值域.



解　a. 计算

$$T(\boldsymbol{u})=A\boldsymbol{u}=\begin{bmatrix}1&-3\\3&5\\-1&7\end{bmatrix}\begin{bmatrix}2\\-1\end{bmatrix}=\begin{bmatrix}5\\1\\-9\end{bmatrix}$$

b. 解 $T(\boldsymbol{x})=\boldsymbol{b}$，即解 $A\boldsymbol{x}=\boldsymbol{b}$，或

$$\begin{bmatrix}1&-3\\3&5\\-1&7\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}3\\2\\-5\end{bmatrix}\tag{1}$$

应用 1.4 节的方法，把增广矩阵进行行变换：

$$\begin{bmatrix}1&-3&3\\3&5&2\\-1&7&-5\end{bmatrix}\sim\begin{bmatrix}1&-3&3\\0&14&-7\\0&4&-2\end{bmatrix}\sim\begin{bmatrix}1&-3&3\\0&1&-0.5\\0&0&0\end{bmatrix}\sim\begin{bmatrix}1&0&1.5\\0&1&-0.5\\0&0&0\end{bmatrix}\tag{2}$$

因此 $x_1=1.5, x_2=-0.5, \boldsymbol{x}=\begin{bmatrix}1.5\\-0.5\end{bmatrix}$，这个向量 $\boldsymbol{x}$ 在 T 下的像是给定的向量 $\boldsymbol{b}$.

c. 任意 $\boldsymbol{x}$，若它在 T 下的像是 $\boldsymbol{b}$，它必满足（1），由（2）知方程（1）的解是唯一的，所以仅有一个 $\boldsymbol{x}$ 使它的像是 $\boldsymbol{b}$.

d. 若 $\boldsymbol{c}$ 是 $\mathbb{R}^2$ 中某个 $\boldsymbol{x}$ 在 T 下的像，则它属于 T 的值域，也就是说，对某个 $\boldsymbol{x}$，$\boldsymbol{c}=T(\boldsymbol{x})$，这就是说，要问方程组 $A\boldsymbol{x}=\boldsymbol{c}$ 是否相容，为找出答案，把增广矩阵进行行变换：

$$\begin{bmatrix}1 & -3 & 3\\ 3 & 5 & 2\\ -1 & 7 & 5\end{bmatrix} \sim \begin{bmatrix}1 & -3 & 3\\ 0 & 14 & -7\\ 0 & 4 & 8\end{bmatrix} \sim \begin{bmatrix}1 & -3 & 3\\ 0 & 1 & 2\\ 0 & 14 & -7\end{bmatrix} \sim \begin{bmatrix}1 & -3 & 3\\ 0 & 1 & 2\\ 0 & 0 & -35\end{bmatrix}$$

第 3 个方程是 $0=-35$，说明方程组不相容，因此 $\boldsymbol{c}$ 不属于 T 的值域. ■

例 1（c）的问题是方程组的唯一性问题：$\boldsymbol{b}$ 是否 $\mathbb{R}^n$ 中唯一的 $\boldsymbol{x}$ 的像？类似地，例 1（d）是存在性问题：是否存在 $\mathbb{R}^n$ 中 $\boldsymbol{x}$ 使它的像为 $\boldsymbol{c}$？

下面两个矩阵变换有很明确的几何意义，它们加强了矩阵作为向量变换的动态观点，2.7 节包含了其他有关计算机图形学的有趣例子.

例 2 若 $A=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{bmatrix}$，则变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 把 $\mathbb{R}^3$ 中的点投影到 x_1x_2 坐标平面上，因为

$$\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} \mapsto \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} = \begin{bmatrix}x_1\\ x_2\\ 0\end{bmatrix}$$

见图 1-36.

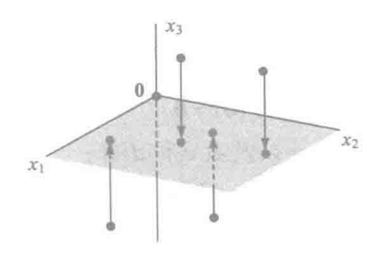
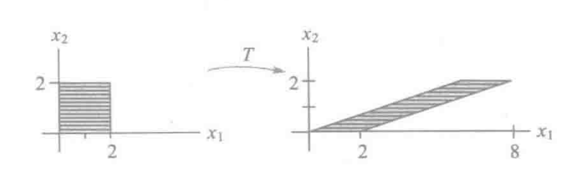


图 1-36 投影变换 ■

例 3 设 $A=\begin{bmatrix}1 & 3\\ 0 & 1\end{bmatrix}$，变换 $T:\mathbb{R}^2\to\mathbb{R}^2$ 定义为 $T(\boldsymbol{x})=A\boldsymbol{x}$ 称为**剪切变换**. 可以说明，若 T 作用于图 1-37 的 2×2 正方形的各点，则像的集构成带阴影的平行四边形. 关键的想法是证明 T 将线段映射成为线段（如习题 27 所示），然后验证正方形的 4 个顶点映射成平行四边形的 4 个顶点. 例如，点 $\boldsymbol{u}=\begin{bmatrix}0\\ 2\end{bmatrix}$ 的像为 $T(\boldsymbol{u})=\begin{bmatrix}1 & 3\\ 0 & 1\end{bmatrix}\begin{bmatrix}0\\ 2\end{bmatrix}=\begin{bmatrix}6\\ 2\end{bmatrix}$，$\begin{bmatrix}2\\ 2\end{bmatrix}$ 的像为 $\begin{bmatrix}1 & 3\\ 0 & 1\end{bmatrix}\begin{bmatrix}2\\ 2\end{bmatrix}=\begin{bmatrix}8\\ 2\end{bmatrix}$. T 将正方形变形，正方形的底保持不变，而正方形的顶拉向右边. 剪切变换出现在物理学、地质学与晶体学中.



绵羊

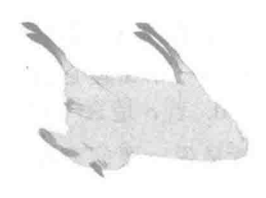

剪切变换后的绵羊

图 1-37 剪切变换 ■

线性变换

1.4 节定理 5 表明，若 A 是 $m\times n$ 矩阵，则变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 有以下性质：

$$A(\boldsymbol{u}+\boldsymbol{v})=A\boldsymbol{u}+A\boldsymbol{v}\text{和}A(c\boldsymbol{u})=cA\boldsymbol{u}$$

$\boldsymbol{u},\boldsymbol{v}$ 是 $\mathbb{R}^n$ 中任意向量，c 是任意数. 这些性质若用函数记号来表示，就得到线性代数中最重要的一类变换.

定义 变换（或映射）T 称为**线性的**，若

（i）对 T 的定义域中一切 $\boldsymbol{u},\boldsymbol{v}$，$T(\boldsymbol{u}+\boldsymbol{v})=T(\boldsymbol{u})+T(\boldsymbol{v})$.

（ii）对一切 $\boldsymbol{u}$ 和标量 c，$T(c\boldsymbol{u})=cT(\boldsymbol{u})$.

每个矩阵变换都是线性变换，非矩阵变换的线性变换的重要例子将在第 4、5 章中讨论.

线性变换保持向量的加法运算与标量乘法运算. 性质（i）说明先将 $\mathbb{R}^n$ 中的 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 相加然后再作用以 T 的结果 $T(\boldsymbol{u}+\boldsymbol{v})$ 等于先把 T 作用于 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 然后将 $\mathbb{R}^m$ 中的 $T(\boldsymbol{u})$ 和 $T(\boldsymbol{v})$ 相加. 由这两个性质容易推出下列性质：

> 若 T 是线性变换，则
>
> $$T(\boldsymbol{0})=\boldsymbol{0} \qquad (3)$$
>
> 且对 T 的定义域中一切向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 以及数 c 和 d 有：
>
> $$T(c\boldsymbol{u}+d\boldsymbol{v})=cT(\boldsymbol{u})+dT(\boldsymbol{v}) \qquad (4)$$

性质（3）由（ii）得出，因 $T(\boldsymbol{0})=T(0\boldsymbol{u})=0T(\boldsymbol{u})=\boldsymbol{0}$.性质（4）由（i）和（ii）推出：

$$T(c\boldsymbol{u}+d\boldsymbol{v})=T(c\boldsymbol{u})+T(d\boldsymbol{v})=cT(\boldsymbol{u})+dT(\boldsymbol{v})$$

若一个变换满足（4），它必是线性的.（取 $c=d=1$ 可得（i），取 $d=0$ 可得（ii）.） 重复应用（4）得出有用的推广：

$$T(c_1\boldsymbol{v}_1+\cdots+c_p\boldsymbol{v}_p)=c_1T(\boldsymbol{v}_1)+\cdots+c_pT(\boldsymbol{v}_p) \qquad (5)$$

在工程和物理中，（5）式称为*叠加原理*. 设想 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p$ 为进入某个系统的信号，$T(\boldsymbol{v}_1),\cdots,T(\boldsymbol{v}_p)$ 为系统对这些信号的响应. 系统满足叠加原理，若某一输入可表示为这些信号的线性组合，则系统的响应是对各个信号的响应的同样的线性组合. 我们将在第 4 章研究这一思想.

例 4 给定标量 r，定义 $T:\mathbb{R}^2\to\mathbb{R}^2$ 为 $T(\boldsymbol{x})=r\boldsymbol{x}$. 当 $0\leqslant r\leqslant 1$ 时，T 称为**压缩变换**，当 $r>1$

时 T 称为**拉伸变换**. 设 $r=3$，证明 T 是线性变换.

解 设 $\boldsymbol{u},\boldsymbol{v}$ 属于 $\mathbb{R}^2$，c,d 为数，则

$$\begin{aligned} T(c\boldsymbol{u}+d\boldsymbol{v}) &= 3(c\boldsymbol{u}+d\boldsymbol{v}) && T \text{ 的定义} \\ &= 3c\boldsymbol{u}+3d\boldsymbol{v} && \text{向量运算} \\ &= c(3\boldsymbol{u})+d(3\boldsymbol{v}) && \\ &= cT(\boldsymbol{u})+dT(\boldsymbol{v}) \end{aligned}$$

因满足（4），于是 T 是线性变换. 见图 1-38.

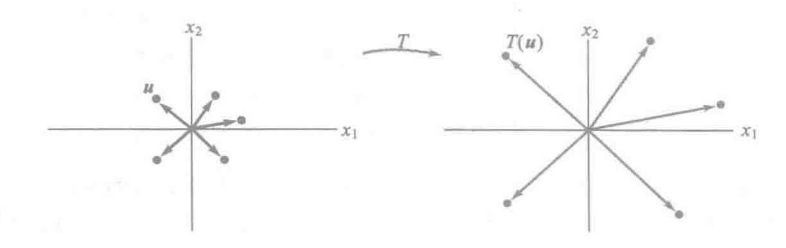


图 1-38 拉伸变换 ■

例 5 定义线性变换 $T:\mathbb{R}^2 \to \mathbb{R}^2$ 为

$$T(\boldsymbol{x})=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}-x_2\\ x_1\end{bmatrix}$$

求出 $\boldsymbol{u}=\begin{bmatrix}4\\ 1\end{bmatrix},\boldsymbol{v}=\begin{bmatrix}2\\ 3\end{bmatrix}$和$\boldsymbol{u}+\boldsymbol{v}=\begin{bmatrix}6\\ 4\end{bmatrix}$在 T 下的像.

解

$$T(\boldsymbol{u})=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}4\\ 1\end{bmatrix}=\begin{bmatrix}-1\\ 4\end{bmatrix},T(\boldsymbol{v})=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}2\\ 3\end{bmatrix}=\begin{bmatrix}-3\\ 2\end{bmatrix},\ T(\boldsymbol{u}+\boldsymbol{v})=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}6\\ 4\end{bmatrix}=\begin{bmatrix}-4\\ 6\end{bmatrix}$$

注意 $T(\boldsymbol{u}+\boldsymbol{v})$ 等于 $T(\boldsymbol{u})+T(\boldsymbol{v})$. 由图 1-39 可知，$T$ 把 $\boldsymbol{u},\boldsymbol{v}$ 和 $\boldsymbol{u}+\boldsymbol{v}$ 逆时针向旋转 90°. 事实上，T 把由 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 确定的平行四边形变换成由 $T(\boldsymbol{u}),T(\boldsymbol{v})$ 确定的平行四边形（见习题 28）.

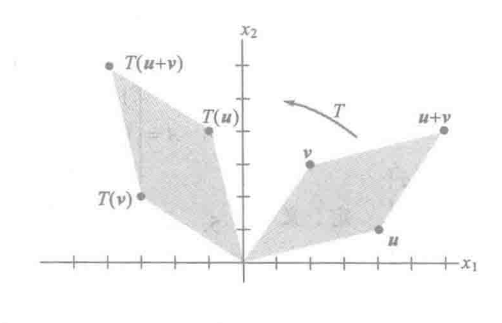


图 1-39 旋转变换 ■

最后的例子不是几何的，它说明线性变换如何把某一类型的数据变成另一种类型的数据.

例 6 某公司生产两种产品 B 和 C. 使用 1.3 节例 7 中的数据，我们构造"单位成本"矩阵 $U=[\boldsymbol{b}\ \ \boldsymbol{c}]$，它的各列描述"每美元产出成本"：

$$\begin{array}{c} \text{产}\quad\ \text{品} \\ \text{B}\qquad \text{C} \end{array}$$

$$U=\begin{bmatrix}0.45 & 0.40\\ 0.25 & 0.35\\ 0.15 & 0.15\end{bmatrix}\begin{matrix}\text{材料}\\ \text{劳动}\\ \text{管理}\end{matrix}$$

设 $\boldsymbol{x}=(x_1,x_2)$ 为"产出"向量，对应于产品 B 和 x_1 美元及产品 C 和 x_2 美元，定义 $T:\mathbb{R}^2\to\mathbb{R}^3$ 为

$$T(\boldsymbol{x})=U\boldsymbol{x}=x_1\begin{bmatrix}0.45\\0.25\\0.15\end{bmatrix}+x_2\begin{bmatrix}0.40\\0.35\\0.15\end{bmatrix}=\begin{bmatrix}\text{材料总成本}\\ \text{劳动总成本}\\ \text{管理总成本}\end{bmatrix}$$

变换 T 将一列产出数量（以美元计算）变换为一列总成本. 该变换是线性的，体现在两方面：

1. 若产量增加 4 倍，由 $\boldsymbol{x}$ 增加到 $4\boldsymbol{x}$，则成本也乘以同一因子，由 $T(\boldsymbol{x})$ 增加到 $4T(\boldsymbol{x})$.

2. 若 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 为产出向量，则对应产出向量 $\boldsymbol{x}+\boldsymbol{y}$ 的总成本恰好等于成本向量 $T(\boldsymbol{x})$ 和 $T(\boldsymbol{y})$ 之和. ■

练习题

1. 设 $T:\mathbb{R}^5\to\mathbb{R}^2, T(\boldsymbol{x})=A\boldsymbol{x}, A$ 为某个矩阵，$\boldsymbol{x}$ 属于 $\mathbb{R}^5$，A 应有几行几列？
2. 设 $A=\begin{bmatrix}1&0\\0&-1\end{bmatrix}$，给出变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的几何解释.
3. 由 $\boldsymbol{0}$ 到向量 $\boldsymbol{u}$ 的线段是形如 $t\boldsymbol{u}$ 的点的集合，其中 $0\leqslant t\leqslant 1$，证明线性变换 T 把这个线段变为 $\boldsymbol{0}$ 到 $T(\boldsymbol{u})$ 的线段.

习题 1.8

1. 设 $A=\begin{bmatrix}2&0\\0&2\end{bmatrix}$，定义 $T:\mathbb{R}^2\to\mathbb{R}^2$ 为 $T(\boldsymbol{x})=A\boldsymbol{x}$，求出 $\boldsymbol{u}=\begin{bmatrix}1\\-3\end{bmatrix}$ 与 $\boldsymbol{v}=\begin{bmatrix}a\\b\end{bmatrix}$ 在 T 下的像.

2. 设 $A=\begin{bmatrix}\frac{1}{3}&0&0\\0&\frac{1}{3}&0\\0&0&\frac{1}{3}\end{bmatrix}, \boldsymbol{u}=\begin{bmatrix}3\\6\\-9\end{bmatrix}, \boldsymbol{v}=\begin{bmatrix}a\\b\\c\end{bmatrix}$，定义 $T:\mathbb{R}^3\to\mathbb{R}^3$ 为 $T(\boldsymbol{x})=A\boldsymbol{x}$，求 $T(\boldsymbol{u})$ 和 $T(\boldsymbol{v})$.

习题 3~6 中，定义 $T(\boldsymbol{x})=A\boldsymbol{x}$，求出 $\boldsymbol{x}$ 使它在 T 下的像为 $\boldsymbol{b}$，并判断 $\boldsymbol{x}$ 是否唯一.

3. $A=\begin{bmatrix}1&0&-3\\-3&1&6\\2&-2&-1\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}-2\\3\\-1\end{bmatrix}$

4. $A=\begin{bmatrix}1&-2&3\\0&1&-3\\2&-5&6\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}-6\\-4\\-5\end{bmatrix}$

5. $A=\begin{bmatrix}1&-5&-7\\-3&7&5\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}-2\\-2\end{bmatrix}$

6. $A=\begin{bmatrix}1&-3&2\\3&-8&8\\0&1&2\\1&0&8\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}1\\6\\3\\10\end{bmatrix}$

7. 设 A 是 6×5 矩阵，为了定义 $T:\mathbb{R}^a\to\mathbb{R}^b$，$T(\boldsymbol{x})=A\boldsymbol{x}$，$a$ 与 b 应为多少？

8. 为了定义从 $\mathbb{R}^5$ 到 $\mathbb{R}^7$ 的映射 $T(\boldsymbol{x})=A\boldsymbol{x}$，矩阵 A 应有几行几列？

习题 9~10 中，求出 $\mathbb{R}^4$ 中所有 $\boldsymbol{x}$，它在变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 下映射为零向量.

9. $A=\begin{bmatrix}1&-3&5&-5\\0&1&-3&5\\2&-4&4&-4\end{bmatrix}$ 10. $A=\begin{bmatrix}3&2&10&-6\\1&0&2&-4\\0&1&2&3\\1&4&10&8\end{bmatrix}$

11. 设 $\boldsymbol{b}=\begin{bmatrix}-1\\1\\0\end{bmatrix}$，$A$ 为习题 9 中的矩阵，$\boldsymbol{b}$ 是否属于线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的值域？

12. 设 $\boldsymbol{b}=\begin{bmatrix}-1\\3\\-1\\4\end{bmatrix}$，$A$ 为习题 10 中的矩阵，$\boldsymbol{b}$ 是否属于线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的值域？

习题 13~16 中，在直角坐标系下标出向量 $\boldsymbol{u}=\begin{bmatrix}5\\2\end{bmatrix},\boldsymbol{v}=\begin{bmatrix}-2\\4\end{bmatrix}$ 和它们在变换 T 下的像(每题画一个图)，给出变换 T 对 $\mathbb{R}^2$ 中向量 $\boldsymbol{x}$ 的作用.

13. $T(\boldsymbol{x})=\begin{bmatrix}-1&0\\0&-1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$

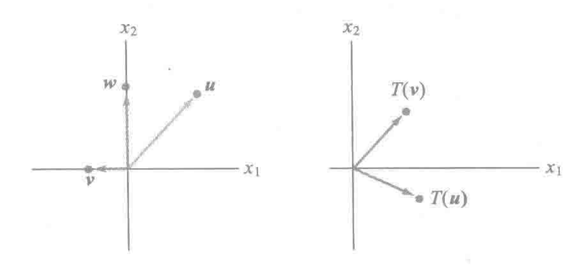
14. $T(\boldsymbol{x})=\begin{bmatrix}2&0\\0&2\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$

15. $T(\boldsymbol{x})=\begin{bmatrix}0&1\\1&0\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$

16. $T(\boldsymbol{x})=\begin{bmatrix}0&0\\0&2\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$

17. 设 $T:\mathbb{R}^2\to\mathbb{R}^2$ 是线性变换，把 $\boldsymbol{u}=\begin{bmatrix}3\\4\end{bmatrix}$ 变为 $\begin{bmatrix}4\\1\end{bmatrix}$，把 $\boldsymbol{v}=\begin{bmatrix}3\\3\end{bmatrix}$ 变为 $\begin{bmatrix}-1\\3\end{bmatrix}$，利用 T 是线性变换的事实求出向量 $2\boldsymbol{u}$, $3\boldsymbol{v}$, $2\boldsymbol{u}+3\boldsymbol{v}$ 在 T 下的像.

18. 下图给出向量 $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}$，以及在线性变换 $T:\mathbb{R}^2\to\mathbb{R}^2$ 的作用下 $T(\boldsymbol{u})$ 和 $T(\boldsymbol{v})$ 的像. 画出 $T(\boldsymbol{w})$ 的像.（提示：首先将 $\boldsymbol{w}$ 写成 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的线性组合.）



19. 设 $\boldsymbol{e}_1=\begin{bmatrix}1\\0\end{bmatrix},\boldsymbol{e}_2=\begin{bmatrix}0\\1\end{bmatrix},\boldsymbol{y}_1=\begin{bmatrix}2\\5\end{bmatrix},\boldsymbol{y}_2=\begin{bmatrix}-1\\6\end{bmatrix},T:\mathbb{R}^2\to\mathbb{R}^2$ 是线性变换，把 $\boldsymbol{e}_1$ 变为 $\boldsymbol{y}_1$，把 $\boldsymbol{e}_2$ 变为 $\boldsymbol{y}_2$，求 $\begin{bmatrix}5\\-3\end{bmatrix}$ 和 $\begin{bmatrix}x_1\\x_2\end{bmatrix}$ 的像.

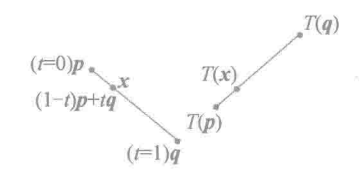
20. 设 $\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\end{bmatrix},\boldsymbol{v}_1=\begin{bmatrix}-3\\5\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}7\\-2\end{bmatrix},T:\mathbb{R}^2\to\mathbb{R}^2$ 是线性变换，把 $\boldsymbol{x}$ 映射为 $x_1\boldsymbol{v}_1+x_2\boldsymbol{v}_2$，求矩阵 A，使对每个 $\boldsymbol{x}$，$T(\boldsymbol{x})=A\boldsymbol{x}$.

习题 21~22 中，判断每个命题的真假，并说明理由.

21. a. 线性变换是一种特殊的函数.
    b. 若 A 是 3×5 矩阵，T 是由 $T(\boldsymbol{x})=A\boldsymbol{x}$ 定义的变换，则 A 的定义域是 $\mathbb{R}^3$.
    c. 若 A 是 $m\times n$ 矩阵，则线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的值域是 $\mathbb{R}^m$.
    d. 所有线性变换都是矩阵变换.
    e. 变换 T 是线性的，当且仅当对任意 T 的定义域中的 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 和所有数 c_1,c_2 都有
    $$T(c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2)=c_1T(\boldsymbol{v}_1)+c_2T(\boldsymbol{v}_2)$$

22. a. 变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的值域是 A 的列的所有线性组合所构成的集合.
    b. 所有矩阵变换都是线性变换.
    c. 若 $T:\mathbb{R}^n\to\mathbb{R}^m$ 是线性变换，c 在 $\mathbb{R}^m$ 中，则唯一性问题是：“c 是否在 T 的值域中？”
    d. 线性变换保持向量加法和标量乘法运算.
    e. 线性变换 $T:\mathbb{R}^n\to\mathbb{R}^m$ 总是把 $\mathbb{R}^n$ 的原点变为 $\mathbb{R}^m$ 的原点.

23. 定义 $f:\mathbb{R}\to\mathbb{R}$为$f(x)=mx+b$.
  a. 证明当 $b=0$ 时，f 是线性变换.
  b. 求出线性变换的一个性质，使当 $b\neq 0$ 时，f 不满足.
  c. 为什么称 f 为线性函数？
24. 仿射变换 $T:\mathbb{R}^n\to\mathbb{R}^m$ 有形式 $T(\boldsymbol{x})=A\boldsymbol{x}+\boldsymbol{b}$, A 为 $m\times n$ 矩阵，$\boldsymbol{b}$ 属于 $\mathbb{R}^m$，证明当 $\boldsymbol{b}\neq\boldsymbol{0}$ 时，T 不是线性变换(仿射变换在计算机图形学中很重要).
25. 给定 $\mathbb{R}^n$ 中向量 $\boldsymbol{v}\neq\boldsymbol{0}$ 和 $\boldsymbol{p}$，通过 $\boldsymbol{p}$ 方向为 $\boldsymbol{v}$ 的直线有参数方程 $\boldsymbol{x}=\boldsymbol{p}+t\boldsymbol{v}$. 证明线性变换 $T:\mathbb{R}^n\to\mathbb{R}^m$ 把此直线映射到另一条直线或一点（称为退化直线）上.
26. a. 证明通过 $\mathbb{R}^n$ 中向量 $\boldsymbol{p}$ 和 $\boldsymbol{q}$ 的直线可写成参数方程 $\boldsymbol{x}=(1-t)\boldsymbol{p}+t\boldsymbol{q}$.（参阅 1.5 节的习题 21 和 22 中的图.）
  b. 由 $\boldsymbol{p}$ 到 $\boldsymbol{q}$ 的线段可表示为所有形如 $(1-t)\boldsymbol{p}+t\boldsymbol{q}$ $(0\leqslant t\leqslant 1)$ 的点的集合（如下图所示）.证明任意一个线性变换 T 把此线段映射为一条线段或一个单独的点.



27. 设 $\boldsymbol{u},\boldsymbol{v}$ 为 $\mathbb{R}^3$ 中线性无关向量，P 是通过 $\boldsymbol{u},\boldsymbol{v}$ 和 $\boldsymbol{0}$ 的平面，P 的参数方程为
$$x=s\boldsymbol{u}+t\boldsymbol{v}(s,t\text{为实数}).$$
证明任意线性变换 $T:\mathbb{R}^3\to\mathbb{R}^3$ 把 P 映射为通过 $\boldsymbol{0}$ 的一个平面，或通过 $\boldsymbol{0}$ 的一条直线，或仅是 $\mathbb{R}^3$ 中的原点.为了使平面 P 的像是平面，$T(\boldsymbol{u})$ 和 $T(\boldsymbol{v})$ 应满足什么条件？
28. 设 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是 $\mathbb{R}^n$ 中的向量，可以证明，由 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 所确定的平行四边形内所有点的集 P 可表示为 $a\boldsymbol{u}+b\boldsymbol{v}$，$0\leqslant a\leqslant 1$, $0\leqslant b\leqslant 1$. 设 $T:\mathbb{R}^n\to\mathbb{R}^m$ 为线性变换，说明为什么 P 内一点在 T 下的像是在由 $T(\boldsymbol{u})$ 和 $T(\boldsymbol{v})$ 确定的平行四边形内.
29. 设 T 是把点映射为它关于 x_2 轴的对称点的线性变换 $T:\mathbb{R}^2\to\mathbb{R}^2$. 画出类似图 1-39 的两个图，说明线性变换的性质（i）和（ii）.
30. 设向量 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 生成 $\mathbb{R}^n$，$T:\mathbb{R}^n\to\mathbb{R}^n$ 是一线性变换，$T(\boldsymbol{v}_i)=0$, $i=1,\cdots,p$，证明 T 是零变换，即若 $\boldsymbol{x}$ 是 $\mathbb{R}^n$ 中的任一向量，则 $T(\boldsymbol{x})=\boldsymbol{0}$.
31. 设 $T:\mathbb{R}^n\to\mathbb{R}^m$ 为线性变换，设 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 为 $\mathbb{R}^n$ 中线性相关集，说明为什么集 $\{T(\boldsymbol{v}_1),T(\boldsymbol{v}_2),T(\boldsymbol{v}_3)\}$ 线性相关.

习题 32~36 中，将列向量写成行的形式，如 $\boldsymbol{x}=(x_1,x_2)$，$T(\boldsymbol{x})=T(x_1,x_2)$.

32. 证明由 $T(x_1,x_2)=(x_1-2|x_2|,x_1-4x_2)$ 定义的变换 T 不是线性的.
33. 证明由 $T(x_1,x_2)=(x_1-3x_2,\ x_1-3,\ 2x_1-5x_2)$ 定义的变换 T 不是线性的.
34. 设 $T:\mathbb{R}^3\to\mathbb{R}^3$ 为一变换，它将向量 $\boldsymbol{x}=(x_1,x_2,x_3)$ 映射为关于平面 $x_3=0$ 对称的点 $T(\boldsymbol{x})=T(x_1,x_2,-x_3)$，证明 T 是一线性变换.(思路参见例 4.)
35. 设 $T:\mathbb{R}^3\to\mathbb{R}^3$ 为一变换，它将向量 $\boldsymbol{x}=(x_1,x_2,x_3)$ 映射到平面 $x_2=0$ 上，即 $T(\boldsymbol{x})=T(x_1,0,x_3)$，证明 T 是一线性变换.
36. 设 $T:\mathbb{R}^n\to\mathbb{R}^n$ 为线性变换.证明:若 T 把两个线性无关向量映射为线性相关集，则方程 $T(\boldsymbol{x})=\boldsymbol{0}$ 有非平凡解.(提示：设 $\mathbb{R}^n$ 中 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 线性无关，但 $T(\boldsymbol{u})$和$T(\boldsymbol{v})$ 线性相关. 则 $c_1T(\boldsymbol{u})+c_2T(\boldsymbol{v})=\boldsymbol{0}$ 对某个不全为零的权 c_1 和 c_2 成立，然后使用这一方程.)

[M]习题 37~38 中，给定矩阵确定了一个线性变换 T. 求出使 $T(\boldsymbol{x})=\boldsymbol{0}$ 的所有 $\boldsymbol{x}$.

37. $\begin{bmatrix} 2 & 3 & 5 & -5 \\ -7 & 7 & 0 & 0 \\ -3 & 4 & 1 & 3 \\ -9 & 3 & -6 & -4 \end{bmatrix}$ 38. $\begin{bmatrix} 3 & 4 & -7 & 0 \\ 5 & -8 & 7 & 4 \\ 6 & -8 & 6 & 4 \\ 9 & -7 & -2 & 0 \end{bmatrix}$

39. [M] 设 $\boldsymbol{b}=\begin{bmatrix}8\\7\\5\\-3\end{bmatrix}$, A 为习题 37 中的矩阵，$\boldsymbol{b}$ 是否属于变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的值域？若如此；求出在变换 T 下的像为 $\boldsymbol{b}$ 的 $\boldsymbol{x}$.

40. [M] 设 $\boldsymbol{b}=\begin{bmatrix}4\\-4\\-4\\-7\end{bmatrix}$，设 A 为习题 38 中的矩阵，$\boldsymbol{b}$ 是否属于变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的值域？若是，求出在变换 T 下的像为 $\boldsymbol{b}$ 的 $\boldsymbol{x}$.

练习题答案

1. A 必须有 5 列，$A\boldsymbol{x}$ 才有定义. A 必须有 2 行，T 的余定义域才能是 $\mathbb{R}^2$.
2. 在坐标系中任意取一点看看会发生什么. 例如（4，1）映射为（4，−1），变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 把点映射为它关于 x 轴（或 x_1 轴）的对称点. 见图 1-40.

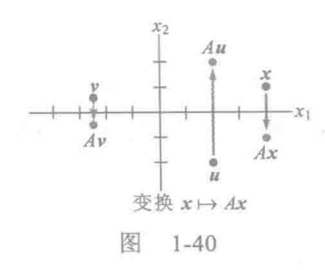


图　1-40

3. 设 $\boldsymbol{x}=t\boldsymbol{u}$, $0\leqslant t\leqslant 1$，因 T 是线性变换，$T(t\boldsymbol{u})=tT(\boldsymbol{u})$，它是连接 $\mathbf{0}$ 和 $T(\boldsymbol{u})$ 的线段上的点.

1.9　线性变换的矩阵

当一个线性变换 T 是由几何中提出来或用语言叙述时，我们通常希望有关于 $T(\boldsymbol{x})$ 的公式. 下面的讨论指出，从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的每一个线性变换，实际上都是一个矩阵变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$，而且变换 T 的性质都归结为 A 的性质. 寻找矩阵 A 的关键，是了解 T 完全由它对单位矩阵 I_n 的各列的作用所决定.

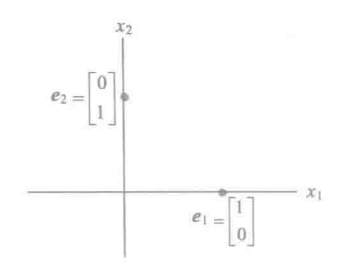
例 1　$I_2=\begin{bmatrix}1&0\\0&1\end{bmatrix}$ 的两列是 $\boldsymbol{e}_1=\begin{bmatrix}1\\0\end{bmatrix}$ 和 $\boldsymbol{e}_2=\begin{bmatrix}0\\1\end{bmatrix}$，设 T 是 $\mathbb{R}^2$ 到 $\mathbb{R}^3$ 的线性变换，使

$$T(\boldsymbol{e}_1)=\begin{bmatrix}5\\-7\\2\end{bmatrix},\ T(\boldsymbol{e}_2)=\begin{bmatrix}-3\\8\\0\end{bmatrix}$$

求出 $\mathbb{R}^2$ 中任意向量 $\boldsymbol{x}$ 的像.

解　写出

$$\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\end{bmatrix}=x_1\begin{bmatrix}1\\0\end{bmatrix}+x_2\begin{bmatrix}0\\1\end{bmatrix}=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2 \tag{1}$$



因为 T 是线性变换，所以

$$T(\boldsymbol{x})=x_1T(\boldsymbol{e}_1)+x_2T(\boldsymbol{e}_2)=x_1\begin{bmatrix}5\\-7\\2\end{bmatrix}+x_2\begin{bmatrix}-3\\8\\0\end{bmatrix}=\begin{bmatrix}5x_1-3x_2\\-7x_1+8x_2\\2x_1+0\end{bmatrix} \tag{2}$$

■

由步骤（1）到（2）说明为什么只要知道 $T(\boldsymbol{e}_1)$ 和 $T(\boldsymbol{e}_2)$ 就可由任意 $\boldsymbol{x}$ 决定 $T(\boldsymbol{x})$，此外，因（2）把 $T(\boldsymbol{x})$ 表示为 $T(\boldsymbol{e}_1)$ 和 $T(\boldsymbol{e}_2)$ 的线性组合，我们可把这些向量作为矩阵 A 的各列，而把（2）式写成

$$T(\boldsymbol{x})=[T(\boldsymbol{e}_1)\ \ T(\boldsymbol{e}_2)]\begin{bmatrix}x_1\\x_2\end{bmatrix}=A\boldsymbol{x}$$

定理 10 设 $T:\mathbb{R}^n\to\mathbb{R}^m$ 为线性变换，则存在唯一的矩阵 A，使

$$T(\boldsymbol{x})=A\boldsymbol{x}，对\ \mathbb{R}^n 中一切 \boldsymbol{x}$$

事实上，A 是 $m\times n$ 矩阵，它的第 j 列是向量 $T(\boldsymbol{e}_j)$，其中 $\boldsymbol{e}_j$ 是单位矩阵 $\boldsymbol{I}_n$ 的第 j 列：

$$A=[T(\boldsymbol{e}_1)\ \cdots\ T(\boldsymbol{e}_n)] \tag{3}$$

证 记 $\boldsymbol{x}=I_n\boldsymbol{x}=[\boldsymbol{e}_1\ \cdots\ \boldsymbol{e}_n]\boldsymbol{x}=x_1\boldsymbol{e}_1+\cdots+x_n\boldsymbol{e}_n$，由于 T 是线性变换，知

$$\begin{aligned}T(\boldsymbol{x})&=T(x_1\boldsymbol{e}_1+\cdots+x_n\boldsymbol{e}_n)=x_1T(\boldsymbol{e}_1)+\cdots+x_nT(\boldsymbol{e}_n)\\&=[T(\boldsymbol{e}_1)\ \cdots\ T(\boldsymbol{e}_n)]\begin{bmatrix}x_1\\\vdots\\x_n\end{bmatrix}=A\boldsymbol{x}\end{aligned}$$

A 的唯一性在习题 33 中研究. ■

（3）中矩阵 A 称为**线性变换 T 的标准矩阵**.

现在我们知道，由 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的每个线性变换都是矩阵变换，反之亦然. 术语线性变换强调映射的性质，而矩阵变换描述这样的映射如何实现. 如下例所示.

例 2 对拉伸变换 $T(\boldsymbol{x})=3\boldsymbol{x}$，求标准矩阵.

解 写出

$$T(\boldsymbol{e}_1)=3\boldsymbol{e}_1=\begin{bmatrix}3\\0\end{bmatrix} \text{ 和 } T(\boldsymbol{e}_2)=3\boldsymbol{e}_2=\begin{bmatrix}0\\3\end{bmatrix}$$

$$A=\begin{bmatrix}3&0\\0&3\end{bmatrix}$$
■

例 3 设 $T:\mathbb{R}^2\to\mathbb{R}^2$ 为把 $\mathbb{R}^2$ 中每一个点逆时针旋转角度 φ 的变换. 我们可以证明这个变换是线性变换（见 1.8 节图 1-39），求出这个变换的标准矩阵.

解 $\begin{bmatrix}1\\0\end{bmatrix}$旋转成为$\begin{bmatrix}\cos\varphi\\\sin\varphi\end{bmatrix}$，$\begin{bmatrix}0\\1\end{bmatrix}$旋转成为$\begin{bmatrix}-\sin\varphi\\\cos\varphi\end{bmatrix}$，见图 1-41.

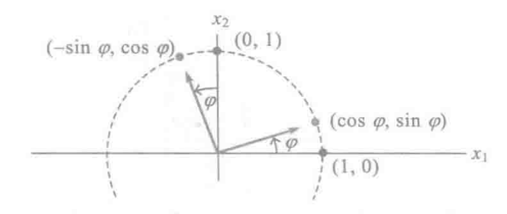


图 1-41 旋转变换

由定理 10，

$$A=\begin{bmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{bmatrix}$$

1.8 节例 5 是这个变换的特殊情形，其中 $\varphi=\pi/2$. ■

$\mathbb{R}^2$ 中的几何线性变换

例 2 和例 3 说明了几何中的线性变换，表 1-2~1-5 说明了其他常见的平面几何线性变换. 因这些变换都是线性的，它们完全由它们对 I_2 的各列的作用确定，而不是仅表示 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$ 的像，下列各表说明了这些变换对单位正方形的作用（见图 1-42）.

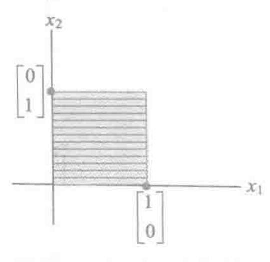


图 1-42 单位正方形

其他的变换可以通过表 1-2~1-5 所列出的变换通过复合构造出来，即一个变换之后再作另一个变换，例如，作一个水平剪切变换后再做一个关于 x_2 轴的对称变换. 2.1 节将证明，线性变换的复合仍是线性的（见习题 34）.

表 1-2 对称

变换	单位正方形的像	标准矩阵
关于 x_1 轴的对称		$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
关于 x_2 轴的对称		$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
关于直线 $x_2=x_1$ 的对称		$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
关于直线 $x_2=-x_1$ 的对称		$\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$
关于原点的对称		$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

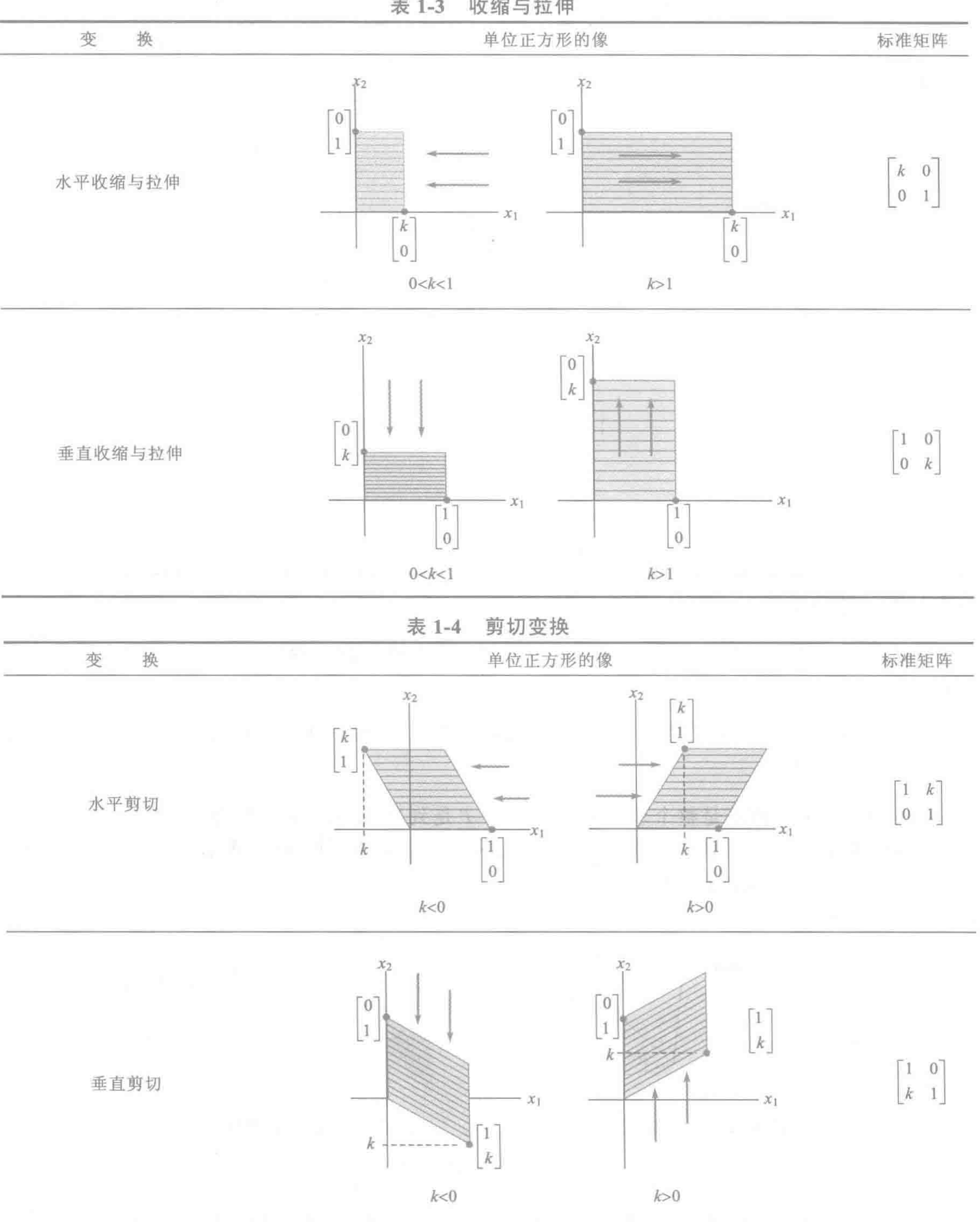


表 1-3　收缩与拉伸

变　换	单位正方形的像	标准矩阵
水平收缩与拉伸	$0<k<1$　　$k>1$	$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$
垂直收缩与拉伸	$0<k<1$　　$k>1$	$\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$

表 1-4　剪切变换

变　换	单位正方形的像	标准矩阵
水平剪切	$k<0$　　$k>0$	$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$
垂直剪切	$k<0$　　$k>0$	$\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$

表 1-5 投影

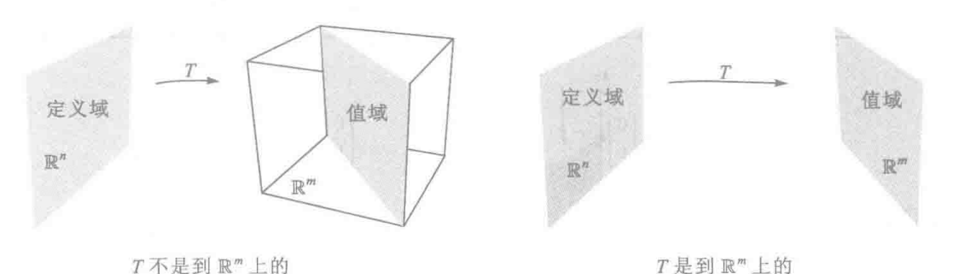
变 换	单位正方形的像	标准矩阵
投影到 x_1 轴上		$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$
投影到 x_2 轴上		$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

存在与唯一性问题

线性变换的概念给出一种新的了解以前提到的存在唯一性问题的观点，下列两个定义给出与变换有关的术语.

定义 映射 $T:\mathbb{R}^n \to \mathbb{R}^m$ 称为到 $\mathbb{R}^m$ 上的映射，若 $\mathbb{R}^m$ 中任一 $\boldsymbol{b}$ 都至少有一个 $\mathbb{R}^n$ 中的 $\boldsymbol{x}$ 与之对应.（也称为**满射**.）

等价地，当 T 的值域是整个余定义域 $\mathbb{R}^m$ 时，T 是到 $\mathbb{R}^m$ 上的. 也就是说，若对 $\mathbb{R}^m$ 中每个 $\boldsymbol{b}$，方程 $T(\boldsymbol{x})=\boldsymbol{b}$ 至少有一个解. “T 是否把 $\mathbb{R}^n$ 映到 $\mathbb{R}^m$ 上？” 是存在性问题. 映射 T 不是到 $\mathbb{R}^m$ 上的，若 $\mathbb{R}^m$ 中有某个 b 使方程 $T(\boldsymbol{x})=\boldsymbol{b}$ 无解. 见图 1-43.



T 不是到 $\mathbb{R}^m$ 上的　　　　T 是到 $\mathbb{R}^m$ 上的

图 1-43 是否 T 的值域是整个 $\mathbb{R}^m$

定义 映射 $T:\mathbb{R}^n \to \mathbb{R}^m$ 称为**一对一**映射（或 1:1），若 $\mathbb{R}^m$ 中每个 b 是 $\mathbb{R}^n$ 中至多一个 x 的像.

（也称为单射.）

等价地，T 是一对一的，若对 $\mathbb{R}^m$ 中每个 $\boldsymbol{b}$，方程 $T(\boldsymbol{x})=\boldsymbol{b}$ 有唯一的解或没有解. “T 是否是一对一的？”是唯一性问题. 映射 T 不是一对一的，若 $\mathbb{R}^m$ 中某个 $\boldsymbol{b}$ 是 $\mathbb{R}^n$ 中多个向量的像. 若没有这样的 $\boldsymbol{b}$，T 就是一对一的. 见图 1-44.

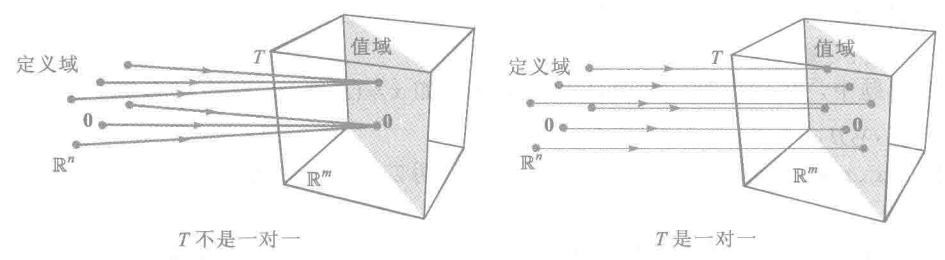


图 1-44　是否每个 $\boldsymbol{b}$ 是至多一个向量的像

表 1-5 中的投影变换不是一对一的，也不能将 $\mathbb{R}^2$ 映上到 $\mathbb{R}^2$. 表 1-2、1-3 和 1-4 中的变换是一对一的，能将 $\mathbb{R}^2$ 映上到 $\mathbb{R}^2$. 其他可能性在下面的两个例子中给出.

例 4 及随后的定理说明了关于一对一映射与映上映射的函数性质是如何与本章以前的一些概念关联起来的.

例 4　设 T 是线性变换，它的标准矩阵为

$$A=\begin{bmatrix}1 & -4 & 8 & 1\\ 0 & 2 & -1 & 3\\ 0 & 0 & 0 & 5\end{bmatrix}$$

T 是否把 $\mathbb{R}^4$ 映上到 $\mathbb{R}^3$？T 是否一对一映射？

解　因 A 已经是阶梯形，可以立即看出，A 在每一行有主元位置，由 1.4 节定理 4，对 $\mathbb{R}^3$ 中每个 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 相容，换句话说，线性变换 T 将 $\mathbb{R}^4$（它的定义域）映射到 $\mathbb{R}^3$ 上. 然而因方程 $A\boldsymbol{x}=\boldsymbol{b}$ 有一个自由变量（因为有 4 个变量，仅有 3 个基本变量），每个 $\boldsymbol{b}$ 都有多个 $\boldsymbol{x}$ 的像，所以 T 不是一对一的. ■

定理 11　设 $T:\mathbb{R}^n\to\mathbb{R}^m$ 为线性变换，则 T 是一对一当且仅当方程 $A\boldsymbol{x}=\boldsymbol{0}$ 仅有平凡解.

证　因 T 是线性的，$T(\boldsymbol{0})=\boldsymbol{0}$，若 T 是一对一的，方程 $T(\boldsymbol{x})=\boldsymbol{0}$ 至多有一个解，因此仅有零解. 若 T 不是一对一的，则 $\mathbb{R}^m$ 中某个 $\boldsymbol{b}$ 是至少 $\mathbb{R}^n$ 中两个相异向量，比如说是 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的像，即 $T(\boldsymbol{u})=\boldsymbol{b},T(\boldsymbol{v})=\boldsymbol{b}$，于是因 T 是线性的.

$$T(\boldsymbol{u}-\boldsymbol{v})=T(\boldsymbol{u})-T(\boldsymbol{v})=\boldsymbol{b}-\boldsymbol{b}=\boldsymbol{0}$$

向量 $\boldsymbol{u}-\boldsymbol{v}$ 不是零，因 $\boldsymbol{u}\neq\boldsymbol{v}$. 因此方程 $T(\boldsymbol{x})=\boldsymbol{0}$ 有多于一个解. 因而定理中两个条件同时成立或同时不成立. ■

定理 12　设 $T:\mathbb{R}^n\to\mathbb{R}^m$ 是线性变换，设 A 为 T 的标准矩阵，则

a. T 把 $\mathbb{R}^n$ 映上到 $\mathbb{R}^m$，当且仅当 A 的列生成 $\mathbb{R}^m$.

b. T 是一对一的，当且仅当 A 的列线性无关.

证 a. 由 1.4 节定理 4，A 的列生成 $\mathbb{R}^m$ 当且仅当方程 $A\boldsymbol{x}=\boldsymbol{b}$ 对每个 $\boldsymbol{b}$ 都相容，换句话说，当且仅当对每个 $\boldsymbol{b}$，方程 $T(\boldsymbol{x})=\boldsymbol{b}$ 至少有一个解，这就是说，T 将 $\mathbb{R}^n$ 映射到 $\mathbb{R}^m$ 上.

b. 方程 $T(\boldsymbol{x})=\boldsymbol{0}$ 和 $A\boldsymbol{x}=\boldsymbol{0}$ 仅有记法不同. 所以由定理 11，T 是一对一的当且仅当 $A\boldsymbol{x}=\boldsymbol{0}$ 仅有平凡解，在 1.7 节（3）的命题已说明，这等价于 A 的各列线性无关. ■

定理 12 的命题（a）等价于命题“T 把 $\mathbb{R}^n$ 映上到 $\mathbb{R}^m$，当且仅当 $\mathbb{R}^m$ 中的任一向量都是 A 的列的一个线性组合.”参见 1.4 节定理 4.

下例以及习题中，我们把列向量写成行的形式，如 $\boldsymbol{x}=(x_1,x_2)$，将 $T(\boldsymbol{x})$ 写成 $T(x_1,x_2)$，以代替更正式的 $T((x_1,x_2))$.

例 5 设 $T(x_1,x_2)=(3x_1+x_2,5x_1+7x_2,x_1+3x_2)$，证明 T 是一对一线性变换. T 是否将 $\mathbb{R}^2$ 映上到 $\mathbb{R}^3$？

解 当 $\boldsymbol{x}$ 和 $T(\boldsymbol{x})$ 写成列向量，容易通过检查 $A\boldsymbol{x}$ 中每个元素的行-向量计算看出 T 的标准矩阵.

$$T(\boldsymbol{x})=\begin{bmatrix}3x_1+\ \ x_2\\5x_1+7x_2\\x_1+3x_2\end{bmatrix}=\underbrace{\begin{bmatrix}?&?\\?&?\\?&?\end{bmatrix}}_{A}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}3&1\\5&7\\1&3\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}\tag{4}$$

故 T 的确是线性变换，它的标准矩阵如（4）所示. A 的列是线性无关的，因为它们互相之间不是倍数关系，由定理 12（b），T 是一对一的. 为确定 T 是否从 $\mathbb{R}^2$ 到 $\mathbb{R}^3$ 的映上映射，观察 A 的各列生成的向量集. 因 A 是 3×2 矩阵，由定理 4 可知，A 的列生成 $\mathbb{R}^3$ 当且仅当 A 有 3 个主元列. 这是不可能的，因 A 仅有 2 列，所以 A 的各列不能生成 $\mathbb{R}^3$，对应的线性变换不是映上到 $\mathbb{R}^3$ 的. 如图 1-45 所示.

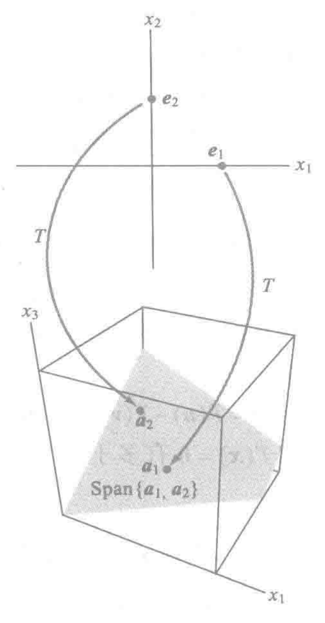


图 1-45 变换 T 不是映上到 $\mathbb{R}^3$ 的 ■

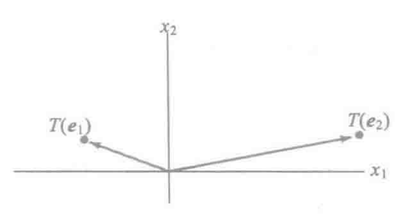
练习题

设 $T:\mathbb{R}^2\to\mathbb{R}^2$ 为一个线性变换，它先作水平剪切变换，将 $\boldsymbol{e}_2$ 映射为 $\boldsymbol{e}_2-0.5\boldsymbol{e}_1$（但 $\boldsymbol{e}_1$ 不变），然后再做关于 x_2 轴的对称变换. 假设 T 是线性的，求它的标准矩阵.（提示：确定 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$ 的像的最终位置.）

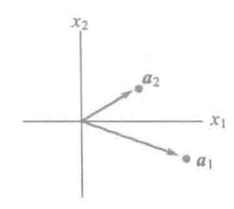
习题 1.9

在习题 1~10 中，设 T 是线性变换，求出 T 的标准矩阵.

1. $T:\mathbb{R}^2\to\mathbb{R}^4$，$T(\boldsymbol{e}_1)=(3,1,3,1)$，$T(\boldsymbol{e}_2)=(-5,2,0,0)$，其中 $\boldsymbol{e}_1=(1,0),\boldsymbol{e}_2=(0,1)$.
2. $T:\mathbb{R}^3\to\mathbb{R}^2$，$T(\boldsymbol{e}_1)=(1,4)$，$T(\boldsymbol{e}_2)=(-2,9)$，$T(\boldsymbol{e}_3)=(3,-8)$，其中 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$ 是 3×3 单位矩阵的列.
3. $T:\mathbb{R}^2\to\mathbb{R}^2$ 是垂直剪切变换，将 $\boldsymbol{e}_1$ 映射为 $\boldsymbol{e}_1-3\boldsymbol{e}_2$ 而保持向量 $\boldsymbol{e}_2$ 不变.
4. $T:\mathbb{R}^2\to\mathbb{R}^2$ 是水平剪切变换，将 $\boldsymbol{e}_2$ 映为 $\boldsymbol{e}_2+2\boldsymbol{e}_1$ 而保持向量 $\boldsymbol{e}_1$ 不变.
5. $T:\mathbb{R}^2\to\mathbb{R}^2$ 将点绕原点逆时针旋转 $\pi/2$ 弧度.
6. $T:\mathbb{R}^2\to\mathbb{R}^2$ 将点绕原点顺时针旋转 $-3\pi/2$ 弧度.
7. $T:\mathbb{R}^2\to\mathbb{R}^2$ 先绕原点顺时针旋转 $-3\pi/4$ 弧度，再关于水平 x_1 轴作对称变换.（提示：$T(\boldsymbol{e}_1)=(-1/\sqrt{2},1/\sqrt{2})$.）
8. $T:\mathbb{R}^2\to\mathbb{R}^2$ 先作水平剪切变换，将 $\boldsymbol{e}_2$ 映为 $\boldsymbol{e}_2+2\boldsymbol{e}_1$ 而保持向量 $\boldsymbol{e}_1$ 不变，再关于直线 $x_2=-x_1$ 作对称变换.
9. $T:\mathbb{R}^2\to\mathbb{R}^2$ 先关于水平 x_1 轴作对称变换，再绕原点顺时针旋转 $-\pi/2$ 弧度.
10. $T:\mathbb{R}^2\to\mathbb{R}^2$ 先关于水平 x_1 轴作对称变换，再关于直线 $x_2=x_1$ 作对称变换.
11. 线性变换 $T:\mathbb{R}^2\to\mathbb{R}^2$ 先关于 x_1 轴作对称变换，再关于 x_2 轴作对称变换. 证明 T 也可被描述为一个绕原点旋转的线性变换. 旋转的角度是多少？
12. 证明：习题 10 中的变换只不过是一个绕原点的旋转. 旋转的角度是多少？
13. 由线性变换 $T:\mathbb{R}^2\to\mathbb{R}^2$ 得到的 $T(\boldsymbol{e}_1)$，$T(\boldsymbol{e}_2)$ 向量如下图所示，画出向量 $T(2,1)$.



14. 线性变换 $T:\mathbb{R}^2\to\mathbb{R}^2$ 的标准矩阵是 $A=[\boldsymbol{a}_1\quad\boldsymbol{a}_2]$，其中 $\boldsymbol{a}_1$和$\boldsymbol{a}_2$ 如下图所示，画出 $\begin{bmatrix}1\\-2\end{bmatrix}$在变换 T 下的像.



习题 15~16 中，填上矩阵中未写出的元素，假设方程对变量的所有值都成立.

15. $$\begin{bmatrix}?&?&?\\?&?&?\\?&?&?\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}2x_1-4x_2\\x_1-x_3\\-x_2+3x_3\end{bmatrix}$$
16. $$\begin{bmatrix}?&?\\?&?\\?&?\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}3x_1-2x_2\\x_1+4x_2\\x_2\end{bmatrix}$$

习题 17~20 中，通过求出映射相应的矩阵，证明 T 是线性变换. 注意 $x_1,x_2,\cdots$ 是向量中的元素而非向量.

17. $T(x_1,x_2,x_3,x_4)=(x_1+2x_2,0,2x_2+x_4,x_2-x_4)$
18. $T(x_1,x_2)=(x_1+4x_2,0,x_1-3x_2,x_1)$
19. $T(x_1,x_2,x_3)=(x_1-5x_2+4x_3,x_2-6x_3)$
20. $T(x_1,x_2,x_3,x_4)=3x_1+4x_3-2x_4\qquad(T:\mathbb{R}^4\to\mathbb{R})$
21. 设 $T:\mathbb{R}^2\to\mathbb{R}^2$ 为使 $T(x_1,x_2)=(x_1+x_2,4x_1+5x_2)$ 的线性变换，求出 $\boldsymbol{x}$，使 $T(\boldsymbol{x})=(3,8)$.

22. 设 $T:\mathbb{R}^2\to\mathbb{R}^3$ 为线性变换，$T(x_1,x_2)=(2x_1-x_2,-3x_1+x_2,2x_1-3x_2)$，求出 $\boldsymbol{x}$，使 $T(\boldsymbol{x})=(0,-1,-4)$.

习题23~24中，判断各个命题的真假，给出理由.

23. a. 线性变换 $T:\mathbb{R}^n\to\mathbb{R}^m$ 完全由它对单位矩阵 I_n 的作用确定.
   b. 若 $T:\mathbb{R}^2\to\mathbb{R}^2$ 把向量绕原点旋转一角度 φ，则 T 是线性变换.
   c. 两个线性变换的复合不一定是线性变换.
   d. 映射 $T:\mathbb{R}^n\to\mathbb{R}^m$ 是把 $\mathbb{R}^n$ 映上到 $\mathbb{R}^m$ 的，若 $\mathbb{R}^n$ 中每个向量 $\boldsymbol{x}$ 映射到 $\mathbb{R}^m$ 中的某个向量.
   e. 若 A 是 3×2 矩阵，则变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 不是一对一的.

24. a. 若 A 是−4×3矩阵，则变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 可以将 $\mathbb{R}^3$ 映上到 $\mathbb{R}^4$.
   b. 每个从 $\mathbb{R}^n$ 映射到 $\mathbb{R}^m$ 中的线性变换都是矩阵变换.
   c. 线性变换 $T:\mathbb{R}^n\to\mathbb{R}^m$ 的标准矩阵的各列就是 $n\times n$ 单位矩阵的各列的像.
   d. 映射 $T:\mathbb{R}^n\to\mathbb{R}^m$ 是一对一的，若 $\mathbb{R}^n$ 中每个向量映射成为 $\mathbb{R}^m$ 中唯一的向量.
   e. 从 $\mathbb{R}^2$ 映射到 $\mathbb{R}^2$ 中的线性变换对点作关于水平轴、垂直轴或原点的对称变换，则其标准矩阵的形式是 $\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$，其中 a和d等于 ±1.

习题 25~28 中，判断给定的线性变换是否是（a）一对一的，（b）满射，给出理由.

25. 习题17中的线性变换.
26. 习题2中的线性变换.
27. 习题19中的线性变换.
28. 习题14中的线性变换.

习题29~30中，使用1.2节例1的符号给出线性变换 T 的标准矩阵可能的阶梯形.

29. $T:\mathbb{R}^3\to\mathbb{R}^4$ 是一对一的.
30. $T:\mathbb{R}^4\to\mathbb{R}^3$ 是满射.
31. 设 $T:\mathbb{R}^n\to\mathbb{R}^m$ 为线性变换，A 为它的标准矩阵，完成下面的命题："T 是一对一的当且仅当 A 有________个主元列". 说明该命题为何是真的.（提示：参看1.7节的习题.）
32. 设 $T:\mathbb{R}^n\to\mathbb{R}^m$ 为线性变换，A 为它的标准矩阵，完成下列命题："T 把 $\mathbb{R}^n$ 映上到 $\mathbb{R}^m$，当且仅当 A 有________个主元列". 根据哪些定理可以确定此命题为真？
33. 证明定理10中 A 的唯一性，设 $T:\mathbb{R}^n\to\mathbb{R}^m$ 为线性变换，对某个 $m\times n$ 矩阵 B，有 $T(\boldsymbol{x})=B\boldsymbol{x}$，证明若 A 是 T 的标准矩阵，则 $A=B$.（提示：证明 A 和 B 有相同的列.）
34. 设 $S:\mathbb{R}^p\to\mathbb{R}^n$，$T:\mathbb{R}^n\to\mathbb{R}^m$ 都是线性变换. 证明映射 $\boldsymbol{x}\mapsto T(S(\boldsymbol{x}))$ 是(从 $\mathbb{R}^p$ 到 $\mathbb{R}^m$ 的)线性映射.（提示：计算 $T(S(c\boldsymbol{u}+d\boldsymbol{v}))$，$\boldsymbol{u},\boldsymbol{v}$ 属于 $\mathbb{R}^p$，c,d 为数. 给出每一步计算的理由，说明为何这一计算给出所需的结论.）
35. 若线性变换 $T:\mathbb{R}^n\to\mathbb{R}^m$ 把 $\mathbb{R}^n$映射到 $\mathbb{R}^m$ 上，你能否给出 m 和 n 之间的一个关系？若 T 是一对一的，你能否给出 m 和 n 之间的关系？
36. 问题"线性变换 T 是否满射？"为什么就是存在性问题？

[M]习题37~40中，设 T 是标准矩阵已给的线性变换，在习题37~38中，判断 T 是否一对一映射，在习题39~40中，判断 T 是否把 $\mathbb{R}^5$ 映上到 $\mathbb{R}^5$，给出理由.

37. $\begin{bmatrix} -5 & 6 & -5 & -6 \\ 8 & 3 & -3 & 8 \\ 2 & 9 & 5 & -12 \\ -3 & 2 & 7 & -12 \end{bmatrix}$ 38. $\begin{bmatrix} 7 & 5 & 9 & -9 \\ 5 & 6 & 4 & -4 \\ 4 & 8 & 0 & 7 \\ -6 & -6 & 6 & 5 \end{bmatrix}$

39. $\begin{bmatrix} 4 & -7 & 3 & 7 & 5 \\ 6 & -8 & 5 & 12 & -8 \\ -7 & 10 & -8 & -9 & 14 \\ 3 & -5 & 4 & 2 & -6 \\ -5 & 6 & -6 & -7 & 3 \end{bmatrix}$

40. $\begin{bmatrix} 9 & 43 & 5 & 6 & -1 \\ 14 & 15 & -7 & -5 & 4 \\ -8 & -6 & 12 & -5 & -9 \\ -5 & -6 & -4 & 9 & 8 \\ 13 & 14 & 15 & 3 & 11 \end{bmatrix}$

练习题答案

看看 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$ 变成什么. 见图 1-46. 首先，$\boldsymbol{e}_1$ 不受剪切变换的影响，而后它被对称变换变为 $-\boldsymbol{e}_1$，所以 $T(\boldsymbol{e}_1)=-\boldsymbol{e}_1$. 其次，$\boldsymbol{e}_2$ 被剪切变换变为 $\boldsymbol{e}_2-0.5\boldsymbol{e}_1$，因关于 x_2 轴的对称变换把 $\boldsymbol{e}_1$ 变为 $-\boldsymbol{e}_1, \boldsymbol{e}_2$ 不变，所以向量 $\boldsymbol{e}_2-0.5\boldsymbol{e}_1$ 变为 $\boldsymbol{e}_2+0.5\boldsymbol{e}_1$，所以 $T(\boldsymbol{e}_2)=\boldsymbol{e}_2+0.5\boldsymbol{e}_1$，于是 T 的标准矩阵为

$$[T(\boldsymbol{e}_1)\ \ T(\boldsymbol{e}_2)]=[-\boldsymbol{e}_1\ \ \boldsymbol{e}_2+0.5\boldsymbol{e}_1]=\begin{bmatrix}-1 & 0.5\\ 0 & 1\end{bmatrix}$$

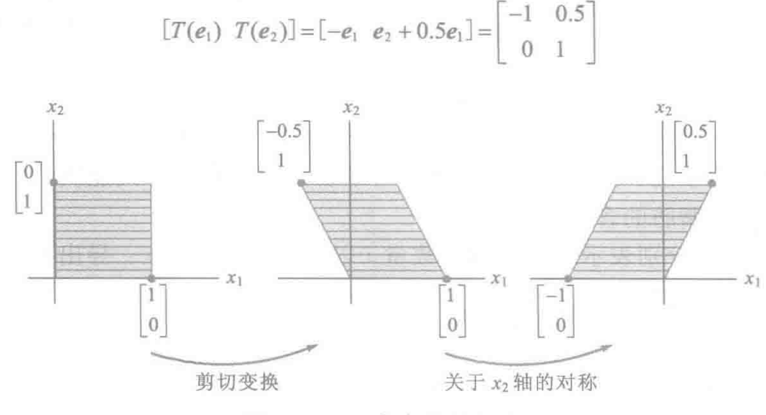


图 1-46　两个变换的复合

1.10　经济学、科学和工程中的线性模型

本节中的数学模型是线性的，也就是说，借助线性方程（通常利用向量或矩阵形式）来描述一个问题，第一个模型讨论营养，实际上代表线性规划问题的一般技术. 第二个模型来自电学. 第三个模型引进线性*差分方程*的概念，此概念是研究动力系统的有力工具，在工程、生态学、通信和管理科学中都有广泛应用. 线性模型的重要性在于当涉及的变量被保持在合理的范围时，自然现象通常是线性或接近线性的. 同时，线性模型比复杂的非线性模型更容易使用计算机来处理.

在阅读每一个模型时，注意线性模型如何体现所研究问题的相关性质.

构造有营养的减肥食谱

一种在 20 世纪 80 年代很流行的食谱，称为剑桥食谱，是经过多年研究编制出来的. 这是由 Alan H. Howard 博士领导的科学家团队经过 8 年对过度肥胖病人的临床研究[⊖]，在剑桥大学完成的. 这种低热量的粉状食品精确地平衡了碳水化合物、高质量的蛋白质和脂肪、配合维生素、矿物质、微量元素和电解质. 近年来，数百万人应用这一食谱实现了快速和有效的减肥.

为得到所希望的数量和比例的营养，Howard 博士在食谱中加入了多种食品. 每种食品供应了多种所需要的成分，然而没有按正确的比例. 例如，脱脂牛奶是蛋白质的主要来源但包含过多的钙，因此大豆粉用来作为蛋白质的来源，它包含较少量的钙. 然而，大豆粉包含过多的脂肪，因而加上乳清，因它含脂肪较少. 然而乳清又含有过多的碳水化合物……

⊖ 这种迅速减肥的方法首先发表在 *International Journal of Obesity*（1978）2, 321-332.

下例说明这个问题小规模的情形. 表 1-6 是该食谱中的 3 种食物以及 100 克每种食物成分含有某些营养素的数量.㊀

表 1-6

营养素（克）	每 100 克成分所含营养素			剑桥食谱每天供应量（克）
	脱脂牛奶	大豆粉	乳清	
蛋白质	36	51	13	33
碳水化合物	52	34	74	45
脂肪	0	7	1.1	3

例 1 求出脱脂牛奶、大豆粉和乳清的某种组合，使该食谱每天能供给表 1-6 中规定的蛋白质、碳水化合物和脂肪的含量.

解 设 x_1, x_2 和 x_3 分别表示这些食物的数量（以 100 克为单位）. 导出方程的一种方法是对每种营养素分别列出方程. 例如，乘积

$$\{x_1 \text{ 单位的脱脂牛奶}\}\{\text{每单位脱脂牛奶所含蛋白质}\}$$

给出 x_1 单位脱脂牛奶供给的蛋白质. 类似地加上大豆粉和乳清所含蛋白质，就应该等于我们所需的蛋白质.类似的计算对每种成分都可进行.

更有效的方法（概念上更为简单）是考虑每种食物的“营养素向量”而建立向量方程. x_1 单位的脱脂牛奶供给的营养素是下列标量乘法:

$$\overset{\text{标量}}{\left\{\begin{matrix} x_1 \text{ 单位的} \\ \text{脱脂牛奶} \end{matrix}\right\}} \cdot \overset{\text{向量}}{\left\{\begin{matrix} \text{每单位脱脂} \\ \text{牛奶的营养素} \end{matrix}\right\}} = x_1\boldsymbol{a}_1 \tag{1}$$

这里 $\boldsymbol{a}_1$ 是表 1-6 的第一列，设 $\boldsymbol{a}_2$ 和 $\boldsymbol{a}_3$ 分别为大豆粉和乳清的对应向量，$\boldsymbol{b}$ 为表示所需要的营养素总量的向量（表中最后一列）. 则 $x_2\boldsymbol{a}_2$ 和 $x_3\boldsymbol{a}_3$ 分别给出由 x_2 单位大豆粉和 x_3 单位乳清给出的营养素，所以所需的方程为

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3 = \boldsymbol{b} \tag{2}$$

把对应的方程组的增广矩阵行变换得

$$\begin{bmatrix} 36 & 51 & 13 & 33 \\ 52 & 34 & 74 & 45 \\ 0 & 7 & 1.1 & 3 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 0 & 0.277 \\ 0 & 1 & 0 & 0.392 \\ 0 & 0 & 1 & 0.233 \end{bmatrix}$$

精确到 3 位小数，该食谱需要 0.277 单位脱脂牛奶、0.392 单位大豆粉、0.233 单位乳清，这样就可供给所需要的蛋白质、碳水化合物与脂肪. ■

重要的是，求出的 x_1, x_2 和 x_3 的值是非负的，这使求出的解有实际意义.（你如何用−0.233 单位乳清？）由于对许多营养素都有要求，可能使用多种食物，以得到有“非负解”的方程组. 因而为了得到这样的解，需要观察各种方程，事实上，剑桥食谱的制造者应用了 33 种食物来供给 31 种营养素.

㊀ 1984 年食谱中的成分；食物成分中的营养素取自 USDA 农业手册，No.8-1 和 8-6，1976.

由食谱构造问题产生线性方程（2），因为由食物供给的营养素可写成一个向量的数量倍，如（1）式所示. 即某种食物供给的营养素与加入到食谱中的此种食物的数量成比例，同时，混合物中的营养素是各种食物中营养素之和.

设计某种特殊的人类或牲畜的食谱问题是经常遇到的，我们构造向量方程的方法常常可以使这些问题的求解得到简化.

线性方程与电路网络

电路网络中的电流可由线性方程组描述. 电源，例如电池，促使电荷在网络中流动. 当电流通过电阻（例如灯泡、电动机等），一部分电压被“用掉”；由欧姆定律，这种“电压降”等于

$$V = RI$$

这里 V 以伏特量度，电阻 R 以欧姆量度（用Ω表示），电流 I 用安培表示（简写为 amps）.

图 1-47 中的网络包含 3 条闭通路，在回路 1，2，3 中流过的电流分别用 I_1, I_2 和 I_3 表示. 回路电流的指定方向是任意取定的，若某一电流求出来是负值，表示实际电流方向与图所选择的方向相反，若所示的电流方向是由电池(-| |-)的正极（长边）指向负极（短边），电压为正，否则电压为负.

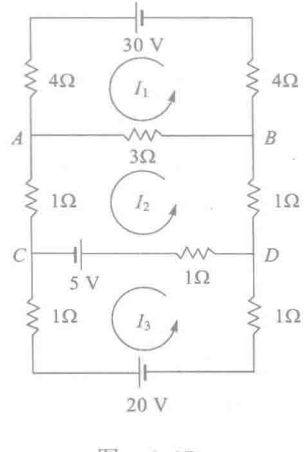


图 1-47

回路中电流服从下列定律.

> **基尔霍夫电压定律**
> 围绕一条回路同一方向的电压降 RI 的代数和等于围绕该回路的同一方向电动势的代数和.

例 2 确定图 1-47 中网络中的回路电流.

解 对回路 1，电流 I_1 通过 3 个电阻，总电压降

$$4I_1 + 4I_1 + 3I_1 = (4+4+3)I_1 = 11I_1$$

回路 2 的电流流过回路 1 的一部分，即 A 与 B 之间的短分支，对应的 RI 电压降为 $3I_2$ 伏特，然而,回路 1 中分支 AB 之间的电流方向与回路 2 中该分支的电流方向相反,因此回路 1 中总的 RI 电压降为 $11I_1-3I_2$.因回路 1 中的电动势为 +30 伏特，基尔霍夫电压定律给出

$$11I_1-3I_2=30$$

对回路 2，方程为

$$-3I_1+6I_2-I_3=5$$

项 $-3I_1$ 来自回路 1 通过分支 AB（电流方向与回路 2 的电流方向相反）的电流，项 $6I_2$ 是回路 2 中所有电阻的和乘以回路电流. 项 $-I_3=-1\cdot I_3$ 是由回路 3 的电流通过 CD 分支 1 欧姆电阻引起的，与回路 2 电流方向相反. 回路 3 的方程为

$$-I_2+3I_3=-25$$

注意分支 CD 上的 5 伏特电池同时属于回路 2 与回路 3，但对回路 3，它是−5 伏特，因它的方向与回路 3 所选择的方向相反.

所以这些电流可由解下列方程组得出

$$\begin{aligned} 11I_1-3I_2 \quad\quad &= 30 \\ -3I_1+6I_2-\ I_3 &= \ 5 \\ -\ I_2+3I_3 &= -25 \end{aligned} \tag{3}$$

对增广矩阵作行变换的解为 $I_1=3$ 安培，$I_2=1$ 安培，$I_3=-8$ 安培，I_3 的负值表示回路 3 的实际方向与图 1-47 中所示相反. ■

把方程组（3）看作向量方程是有启发性的：

$$I_1\begin{bmatrix}11\\-3\\0\end{bmatrix}+I_2\begin{bmatrix}-3\\6\\-1\end{bmatrix}+I_3\begin{bmatrix}0\\-1\\3\end{bmatrix}=\begin{bmatrix}30\\5\\-25\end{bmatrix} \tag{4}$$

$$\quad\uparrow\ \boldsymbol{r}_1 \qquad \uparrow\ \boldsymbol{r}_2 \qquad \uparrow\ \boldsymbol{r}_3 \qquad \uparrow\ \boldsymbol{v}$$

每个向量的第一个元素是在第一个回路中的电阻，类似地第二个、第三个元素分别是在第二、第三个回路中的电阻. 第一个电阻向量 $\boldsymbol{r}_1$ 列出各个回路中电流 I_1 流过的电阻. 当 I_1 流过该电阻的方向与另一回路方向相反时，该电阻取负号. 观察图 1-47，看看如何写出 $\boldsymbol{r}_1$ 中的元素；然后同样写出 $\boldsymbol{r}_2$ 和 $\boldsymbol{r}_3$.（4）的矩阵形式，

$$R\boldsymbol{i}=\boldsymbol{v}, \text{其中 } R=[\boldsymbol{r}_1\ \boldsymbol{r}_2\ \ \boldsymbol{r}_3],\ \boldsymbol{i}=\begin{bmatrix}I_1\\I_2\\I_3\end{bmatrix}$$

给出欧姆定律的矩阵形式. 若所有回路电流都选取同一方向，则 R 的非主对角线元素全部都是负值.

矩阵方程 $R\boldsymbol{i}=\boldsymbol{v}$, 表明这个模型的线性，例如，若电动势向量加倍，则电流向量也加倍；同时，**叠加原理**也成立. 即方程（4）的解是下列方程的解的和

$$Ri=\begin{bmatrix}30\\0\\0\end{bmatrix},Ri=\begin{bmatrix}0\\5\\0\end{bmatrix},\ Ri=\begin{bmatrix}0\\0\\-25\end{bmatrix}$$

这三个方程中的每一个对应回路仅包含一个电动势的网络（其他电源代以封闭该回路的导线）．电流模型是线性的，这是因为欧姆定律与基尔霍夫定律都是线性的；通过某一电阻的电压降与通过它的电流成正比（欧姆定律），而回路中的电压降的和等于回路中电动势的和（基尔霍夫定律）．

网络中回路电流可以决定电路中每一个分支通过的电流．若仅有一回路电流通过该分支，如图 1-47 中从 B 到 D 的分支，则通过该分支的电流等于回路电流．若有多个回路电流通过该分支，如从 A 到 B 的分支，则通过该分支的电流是各回路电流的代数和（基尔霍夫电流定律）．例如，通过分支 AB 的电流为 $I_1-I_2=3-1=2$ 安培，方向与 I_1 相同，分支 CD 中的电流为 $I_2+I_3=9$ 安培．

差分方程

在生态学、经济学和工程技术等领域中，需要研究随时间变化的动力系统，这种系统通常在离散的时刻测量，得到一个向量序列 $\boldsymbol{x}_0,\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots$．向量 $\boldsymbol{x}_k$ 的各个元素给出该系统在第 k 次测量中的状态的信息．

如果有矩阵 A 使 $\boldsymbol{x}_1=A\boldsymbol{x}_0,\boldsymbol{x}_2=A\boldsymbol{x}_1$，一般地，

$$\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k,\quad k=0,1,2,\cdots \tag{5}$$

则（5）式称为**线性差分方程**（或**递归关系**）．给定这样一种关系，我们可由已知的 $\boldsymbol{x}_0$ 计算 $\boldsymbol{x}_1,\boldsymbol{x}_2$ 等等．4.8 节与 4.9 节以及第 5 章的若干节，将推导求 $\boldsymbol{x}_k$ 的公式，并确定 k 无限增大时 $\boldsymbol{x}_k$ 的变化情况．下列的讨论说明导致差分方程问题产生的原因．

地理学家对人口的迁移很有兴趣．这里我们考虑人口在某一城市与它的周边地区之间迁移的简单模型．

固定一个初始年，例如说，2000 年，用 r_0 和 s_0 分别表示该年城市和郊区的人口数．令 $\boldsymbol{x}_0$ 表示人口向量

$$\boldsymbol{x}_0=\begin{bmatrix}r_0\\s_0\end{bmatrix}\begin{matrix}\text{2000年城市人口}\\\text{2000年郊区人口}\end{matrix}$$

对 2001 年与以后各年，把人口向量表示为

$$\boldsymbol{x}_1=\begin{bmatrix}r_1\\s_1\end{bmatrix},\ \boldsymbol{x}_2=\begin{bmatrix}r_2\\s_2\end{bmatrix},\ \boldsymbol{x}_3=\begin{bmatrix}r_3\\s_3\end{bmatrix},\cdots$$

我们的目的是在数学上表示出这些向量的关系．

设人口统计学的研究说明每年约有 5%的城市人口移居郊区（其他 95%留在城市），而 3%的郊区人口移居城市（其他 97%留在郊区）．见图 1-48.

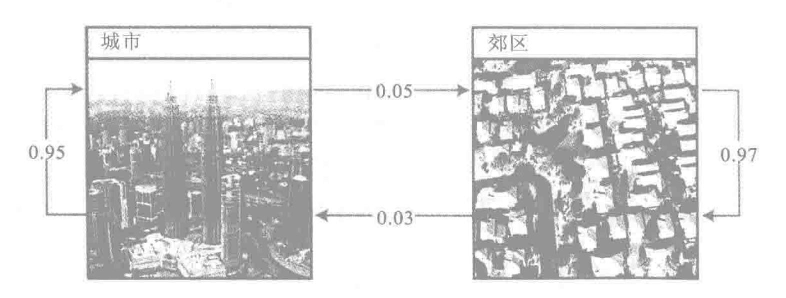


图 1-48 每年城市与郊区迁移的百分比

一年后，原来城市中的人口 r_0 在城市和郊区的分布为

$$\begin{bmatrix} 0.95r_0 \\ 0.05r_0 \end{bmatrix} = r_0 \begin{bmatrix} 0.95 \\ 0.05 \end{bmatrix} \begin{matrix} \text{留在城市} \\ \text{迁到郊区} \end{matrix} \tag{6}$$

郊区 2000 年的人口 s_0 一年后的分配为

$$s_0 \begin{bmatrix} 0.03 \\ 0.97 \end{bmatrix} \begin{matrix} \text{移到城市} \\ \text{留在郊区} \end{matrix} \tag{7}$$

向量（6）和（7）组成 2001 年的全部人口[㊀]，因为

$$\begin{bmatrix} r_1 \\ s_1 \end{bmatrix} = r_0 \begin{bmatrix} 0.95 \\ 0.05 \end{bmatrix} + s_0 \begin{bmatrix} 0.03 \\ 0.97 \end{bmatrix} = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix} \begin{bmatrix} r_0 \\ s_0 \end{bmatrix}$$

即

$$\boldsymbol{x}_1 = M\boldsymbol{x}_0 \tag{8}$$

这里 M 是移民矩阵，由下表确定：

$$\begin{matrix} \text{由：城市} & \text{郊区} & \text{移至：} \\ \end{matrix}$$

$$\begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix} \begin{matrix} \text{城市} \\ \text{郊区} \end{matrix}$$

方程（8）表示人口由 2000 年到 2001 年的变化. 若移民比例保持常数，则由 2000 年到 2001 年的改变为

$$\boldsymbol{x}_2 = M\boldsymbol{x}_1$$

由 2002 年到 2003 年以及以后的各年的变化都是类似的. 一般地

$$\boldsymbol{x}_{k+1} = M\boldsymbol{x}_k,\ k = 0,1,2,\cdots \tag{9}$$

向量序列 $\{\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2, \cdots\}$ 描述了若干年中城市、郊区人口变化的状况.

例 3 设 2000 年城市人口为 600 000，郊区人口为 400 000 人，求上述区域 2001 年和 2002 年的人口.

㊀ 为简单起见，我们忽略出生、死亡、移民等对城市、郊区人口的影响.

解 2000 年的人口为 $\boldsymbol{x}_0 = \begin{bmatrix} 600\,000 \\ 400\,000 \end{bmatrix}$，对 2001 年，

$$\boldsymbol{x}_1 = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix} \begin{bmatrix} 600\,000 \\ 400\,000 \end{bmatrix} = \begin{bmatrix} 582\,000 \\ 418\,000 \end{bmatrix}$$

对 2002 年，

$$\boldsymbol{x}_2 = M\boldsymbol{x}_1 = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix} \begin{bmatrix} 582\,000 \\ 418\,000 \end{bmatrix} = \begin{bmatrix} 565\,440 \\ 434\,560 \end{bmatrix}$$

■

式（9）的人口迁移模型是线性的，因为对应 $\boldsymbol{x}_k \mapsto \boldsymbol{x}_{k+1}$ 是线性变换. 这依赖于两个事实：从一个地区迁往另一个地区的人口与该地区原有的人口成正比，如（6）式和（7）式所示，而这些人口迁移选择的累积效果是不同区域的人口迁移的叠加.

练习题

求出矩阵 A 以及向量 $\boldsymbol{x}$ 和 $\boldsymbol{b}$，使例 1 中的问题成为解方程 $A\boldsymbol{x} = \boldsymbol{b}$.

习题 1.10

1. 一种早餐麦片的包装罐通常列出每份食用量包含的卡路里、蛋白质、碳水化合物与脂肪的量. 两种常见的麦片的营养素含量如下表.

营养素	每份食物营养素含量	
	General Mills Cheerios	Quaker 100%天然麦片
卡路里	110	130
蛋白质（克）	4	3
碳水化合物（克）	20	18
脂肪（克）	2	5

设这两种麦片的混合物要求含热量 295 卡路里、9 克蛋白质、48 克碳水化合物和 8 克脂肪.

a. 建立这个问题的一个向量方程，并给出方程中变量表示的含义.

b. 写出等价的矩阵方程，并判断所希望的两种麦片的混合物是否可以制作出来.

2. 一份脆燕麦片含有 160 卡路里热量、5g 蛋白质、6g 纤维和 1 克脂肪、一份脆片含有 110 卡路里热量、2 克蛋白质、0.1 克纤维和 0.4 克脂肪.

a. 列出矩阵 B 及向量 $\boldsymbol{u}$，使 $B\boldsymbol{u}$ 给出 3 份脆燕麦片和 2 份脆片所含热量、蛋白质、纤维与脂肪的量.

b. [M] 设你希望一种麦片含蛋白质多于脆片但含脂肪少于脆燕麦片. 是否可能混合两种麦片，使它含有 130 卡路里热量、3.20 克蛋白质、2.46 克纤维、0.64 克脂肪？如果可能，如何混合？

3. 上过营养课后，对苹果奶酪情有独钟的 Annie 决定改善午餐，增加西兰花和鸡肉罐头来提高蛋白质和膳食纤维的含量，这道习题的食物营养信息在下表中给出.

每份食物营养素含量				
营养素	苹果奶酪	西兰花	鸡肉罐头	贝类
卡路里	210	51	70	260
蛋白质（克）	10	5.4	15	9
膳食纤维（克）	2	5.0	0	5

a. [M]如果她希望午餐含有 400 卡路里，但要获取 30 克蛋白质、10 克膳食纤维，求苹果奶酪、西兰花和鸡肉罐头之间的比例？

b. [M]她发现（a）中西兰花的比例过高，于是决定用全麦面和白切达奶酪取代传统的苹果奶酪，每种食物的比例为多少能达到（a）一样的结果？

4. 剑桥食谱除例 1 中列出的营养素外，每天还供给 0.8 克钙. 在剑桥食谱中的三种食物，每单位（100 克）供给的钙为：脱脂牛奶 1.26 克、大豆粉 0.19 克、乳清 0.8 克，该食谱中的另一种成分是大豆蛋白质，它每单位供给的营养素为：80 克蛋白质、0 克碳水化合物、3.4 克脂肪和 0.18 克钙.

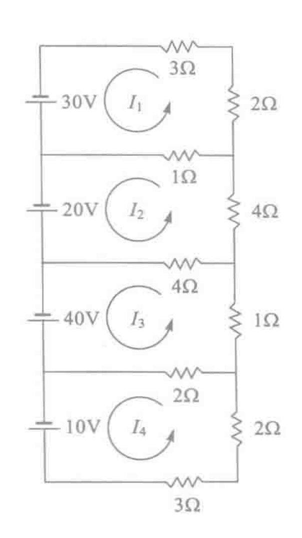
a. 列出矩阵方程，它的解确定脱脂牛奶、大豆粉、乳清与大豆蛋白质的量，使这种混合物正好含剑桥食谱中所需各种营养素的量. 叙述方程中各个变量表示的含义.

b. [M]解（a）中所得方程，讨论你的答案.

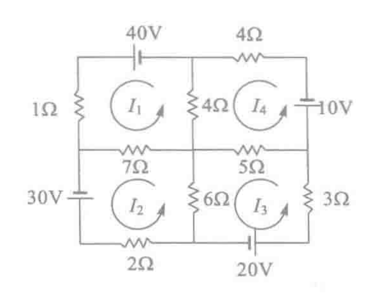
习题 5~8 中，写出矩阵方程以确定回路电流. [M]若能使用 MATLAB 或其他矩阵程序，解出这些方程.

5.

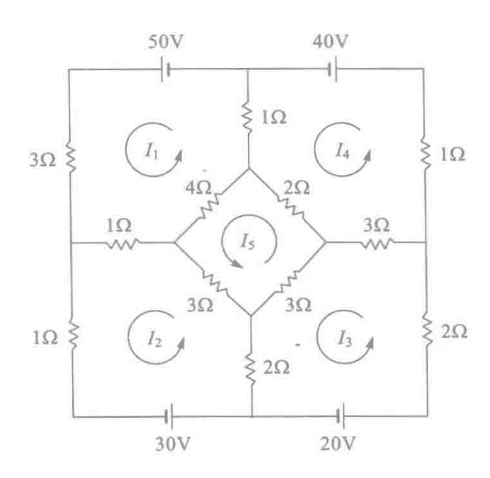
6.



7.



8.



9. 在某区域中，每年有大约7%的城市人口迁往郊区，大约5%的郊区人口迁往城市. 2010年，有800 000 人居住在城市，500 000 人居住在郊区. 列出差分方程描述这些情况，以 $\boldsymbol{x}_0$ 表示2010年的初始人口，然后估计2012年居住在城市和郊区的人口（忽略影响居民人口的其他因素）.

10. 在某区域中，每年大约有 6%的城市人口迁移郊区而4%的郊区居民迁往城市，在2010年，城市有 1 000 000 居民，郊区有 800 000 居民. 列出差分方程描述这些情况，其中 $\boldsymbol{x}_0$ 是 2010 年初始人口. 估计两年后即 2012 年居住在城市和郊区的人口.

11. 在 1994 年初，加利福尼亚的人口为 31 524 000 人，居住在美国其他各州的人口为 228 680 000 人. 该年有 516 100 人由加利福尼亚迁出该州，而381 262人从美国其他地区迁往加利福尼亚. ㊀
    a. 建立该问题的移民矩阵，精确到5位有效数字. 说明如何产生迁移矩阵.
    b. [M]计算 2000 年加利福尼亚和美国其他地区的预计人口数，假设迁移率在6年内不会改变.（不考虑出生、死亡或美国之外的移民.）

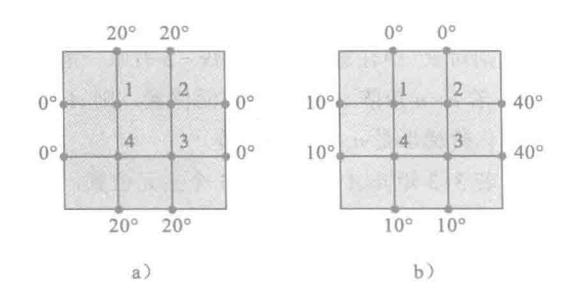
12. [M]堪萨斯州 Wichita 市的 Budget Rent A Car 公司有一个车队共 500 辆车，分布在 3 个地点. 在一个地点租的车可以在3个地点的任何一个交还. 这些车归还的比例在下表中列出. 设星期一有 295 辆车在机场出租，55 辆车在东区出租，150 辆车在西区出租，星期三这些车将会在三个地点如何分配？

从出租地：机场 东区 西区　　归还到：

$$\begin{bmatrix} 0.97 & 0.05 & 0.10 \\ 0.00 & 0.90 & 0.05 \\ 0.03 & 0.05 & 0.85 \end{bmatrix} \begin{matrix} \text{机场} \\ \text{东区} \\ \text{西区} \end{matrix}$$

13. [M]设 M 与 $\boldsymbol{x}_0$ 如例 3.
    a. 计算人口向量 $\boldsymbol{x}_k, k=1,2,\cdots,20$，你发现了什么规律？
    b. 对初始人口为城市 350 000 人，郊区 650 000 人重复（a），你发现了什么规律？

14. [M]研究钢板上边界的温度变化如何影响钢板内部区域的温度.
    a. 首先，估计图中钢板的 4 个点处的温度 T_1,T_2,T_3,T_4，T_k 的值等于最靠近它的 4 个点处温度的平均值.参阅 1.1 节的习题 33 和 34，在那里求得相应的温度值为（20，27.5，30，22.5）（以度为单位），这些数据和你在图 a 和 b 中求出的值有什么关系？
    b. 不做计算，猜想在（a）中边界温度都乘以 3 时钢板内部 4 个点的温度？检验你的猜想.
    c. 最后，猜想边界上 8 个温度与内部 4 个点的温度的相应关系.



a)　　b)

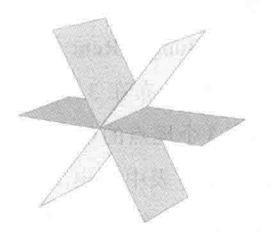
练习题答案

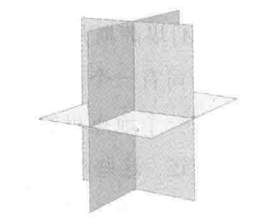
$$A=\begin{bmatrix} 36 & 51 & 13 \\ 52 & 34 & 74 \\ 0 & 7 & 1.1 \end{bmatrix},\ \boldsymbol{x}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},\ \boldsymbol{b}=\begin{bmatrix} 33 \\ 45 \\ 3 \end{bmatrix}$$

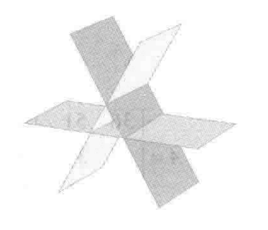
㊀ 由加利福尼亚财政部地理研究所提供的移民数据.

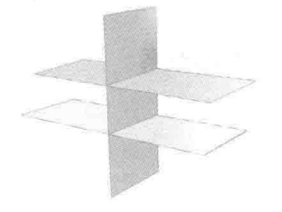
补充习题

1. 标出每个命题的真假，给出理由.（若是真的，举出适当的事实或定理，若是假的，说明原因或举出反例.）
   a. 每个矩阵行等价于唯一的阶梯形矩阵.
   b. 含有 n 个未知数的 n 个方程至多有 n 个解.
   c. 若线性方程组有两个不同的解，则它必有无穷多个解.
   d. 若线性方程组没有自由变量，则它有唯一解.
   e. 若增广矩阵 $[A\ \ \boldsymbol{b}]$ 由初等行变换变为 $[C\ \ \boldsymbol{d}]$，则方程 $A\boldsymbol{x}=\boldsymbol{b}$ 与 $C\boldsymbol{x}=\boldsymbol{d}$ 有相同解集.
   f. 若方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 有多于一个解，则 $A\boldsymbol{x}=\boldsymbol{0}$ 也是.
   g. 若 A 是 $m\times n$ 矩阵，且对某个 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 相容，则 A 的各列生成 $\mathbb{R}^m$.
   h. 若增广矩阵 $[A\ \ \boldsymbol{b}]$ 可由初等行变换化为简化阶梯形，则方程 $A\boldsymbol{x}=\boldsymbol{b}$ 相容.
   i. 若矩阵 A 和 B 行等价，则它们有相同的简化阶梯形.
   j. 方程 $A\boldsymbol{x}=\boldsymbol{0}$ 有平凡解当且仅当它没有自由变量.
   k. 若 A 是 $m\times n$ 矩阵，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 对 $\mathbb{R}^m$ 中任意 $\boldsymbol{b}$ 都相容，则 A 必有 m 个主元列.
   l. 若 $m\times n$ 矩阵 A 在每一行都有一个主元位置，则对 $\mathbb{R}^m$ 中任意 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 有唯一解.
   m. 若 $n\times n$ 矩阵 A 有 n 个主元位置，则 A 的简化阶梯形是 $n\times n$ 单位矩阵.
   n. 若 3×3 矩阵 A 和 B 都有 3 个主元位置，则通过初等行变换可以将 A 变换为 B.
   o. 若 A 是 $m\times n$ 矩阵，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 有至少两个不同的解，如果方程 $A\boldsymbol{x}=\boldsymbol{c}$ 相容，则方程 $A\boldsymbol{x}=\boldsymbol{c}$ 有多个解.
   p. 若 A 和 B 是行等价的 $m\times n$ 矩阵，且 A 的列生成 $\mathbb{R}^m$，则 B 的列也生成 $\mathbb{R}^m$.
   q. 若 R^3 中向量集 $S=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 中任意一个向量都不是其他向量的倍数，则 S 线性无关.
   r. 若 $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$ 是线性无关集，则 $\boldsymbol{u},\boldsymbol{v}$和$\boldsymbol{w}$ 不在 $\mathbb{R}^2$ 中.
   s. 在某些情况下，4 个向量可能生成 $\mathbb{R}^5$.
   t. $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 属于 $\mathbb{R}^m$，则 $-\boldsymbol{u}$ 属于 $\mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$.
   u. 若 $\boldsymbol{u},\boldsymbol{v}$和$\boldsymbol{w}$ 是 $\mathbb{R}^2$ 中的非零向量，则 $\boldsymbol{u}$ 是 $\boldsymbol{u}$和$\boldsymbol{v}$ 的线性组合.
   v. 若 $\boldsymbol{w}$ 是 $\mathbb{R}^n$ 中 $\boldsymbol{u}$和$\boldsymbol{v}$ 的线性组合，则 $\boldsymbol{u}$ 是 $\boldsymbol{v}$和$\boldsymbol{w}$ 的线性组合.
   w. 设 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 是 $\mathbb{R}^5$ 中非零向量，$\boldsymbol{v}_2$ 不是 $\boldsymbol{v}_1$ 的倍数，$\boldsymbol{v}_3$ 不是 $\boldsymbol{v}_1$和$\boldsymbol{v}_2$ 的线性组合，则 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 线性无关.
   x. 线性变换是函数.
   y. 若 A 是 6×5 矩阵，线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 不能将 $\mathbb{R}^5$ 映上到 $\mathbb{R}^6$.
   z. 若 A 是有 m 个主元列的 $m\times n$ 矩阵，则线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 是一对一映射.
2. 设 a 和 b 表示实数. 叙述（线性）方程 $ax=b$ 的解集的各种可能情况.（提示：解的数目依赖于 a 和 b.）
3. 一个线性方程 $ax+by+cz=d$ 的解 (x,y,z) 可以表示为 $\mathbb{R}^3$ 中的一个平面，其中 a,b,c 不全为零. 构造有三个方程的方程组表示图 a 相交于一条直线，b 相交于一点，c 没有交点. 图形如下所示.


a）三个平面交于一条直线


b）三个平面交于一点


c）三个平面没有交点


c′）三个平面没有交点

4. 三个未知量三个方程的线性方程组的系数矩阵的每一列都有一个主元位置，说明方程组为什么有唯一解.

5. 确定 h 和 k 的值，使下列方程组的解集（i）是空集，（ii）包含唯一的解，（iii）包含无穷多个解.

   a. $\begin{aligned} x_1+3x_2&=k \\ 4x_1+hx_2&=8 \end{aligned}$ b. $\begin{aligned} -2x_1+hx_2&=1 \\ 6x_1+kx_2&=-2 \end{aligned}$

6. 确定下列方程组是否相容：

$$\begin{aligned} 4x_1-2x_2+\ \ 7x_3&=-5 \\ 8x_1-3x_2+10x_3&=-3 \end{aligned}$$

   a. 定义适当的向量，把问题重述为线性组合的形式，再解此问题.

   b. 定义适当的矩阵，用“ A 的列”重述此问题.

   c. 定义适当的线性变换 T，利用（b）中矩阵，用 T 的术语重述此问题.

7. 考虑下列的问题，确定方程组是否对任意的 b_1,b_2,b_3 相容.

$$\begin{aligned} 2x_1-4x_2-2x_3&=b_1 \\ -5x_1+\ \ x_2+\ \ x_3&=b_2 \\ 7x_1-5x_2-3x_3&=b_3 \end{aligned}$$

   a. 定义适当的向量，用 $\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 的术语重述问题，然后解此问题.

   b. 定义适当的矩阵 A，用“ A 的列”重述问题.

   c. 用（b）中矩阵定义适当的线性变换 T，用 T 的术语重述问题.

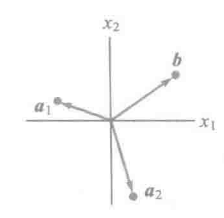
8. 使用 1.2 节例 1 的记号描述矩阵 A 可能的阶梯形.

   a. A 是 2×3 矩阵，其列生成 $\mathbb{R}^2$.

   b. A 是 3×3 矩阵，其列生成 $\mathbb{R}^3$.

9. 将向量 $\begin{bmatrix}5\\6\end{bmatrix}$ 表示为两个向量的和，其中一个在直线 $\{(x,y):y=2x\}$ 上，另一个在直线 $\{(x,y):y=x/2\}$ 上.

10. 设 $\boldsymbol{a}_1,\boldsymbol{a}_2$ 和 $\boldsymbol{b}$ 是 $\mathbb{R}^2$ 中的向量，如下图所示，$A=[\boldsymbol{a}_1\quad\boldsymbol{a}_2]$. 方程 $A\boldsymbol{x}=\boldsymbol{b}$ 是否有解？如果有解，解是否唯一？给出解释.



11. 构造一个 2×3 矩阵 A，不是阶梯形，使得 $A\boldsymbol{x}=\boldsymbol{0}$ 的解是 $\mathbb{R}^3$ 中的一条直线.

12. 构造一个 2×3 矩阵 A，不是阶梯形，使得 $A\boldsymbol{x}=\boldsymbol{0}$ 的解是 $\mathbb{R}^3$ 中的一个平面.

13. 写出一个 3×3 矩阵 A 的阶梯形，使得 A 的首两列是主元列，且 $A\begin{bmatrix}3\\-2\\1\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$.

14. 求 a 的值使得 $\left\{\begin{bmatrix}1\\a\end{bmatrix},\begin{bmatrix}a\\a+2\end{bmatrix}\right\}$ 是线性无关集.

15. 设（a）和（b）中的向量线性无关，数 $a,\cdots,f$ 有何特征？给出理由.（提示：对（b）使用一个定理.）

   a. $\begin{bmatrix}a\\0\\0\end{bmatrix},\begin{bmatrix}b\\c\\0\end{bmatrix},\begin{bmatrix}d\\e\\f\end{bmatrix}$ b. $\begin{bmatrix}a\\1\\0\\0\end{bmatrix},\begin{bmatrix}b\\c\\1\\0\end{bmatrix},\begin{bmatrix}d\\e\\f\\1\end{bmatrix}$

16. 使用 1.7 节定理 7 解释为什么矩阵 A 的列线性无关.

$$A=\begin{bmatrix}1&0&0&0\\2&5&0&0\\3&6&8&0\\4&7&9&10\end{bmatrix}$$

17. 说明为什么 $\mathbb{R}^5$ 中的向量集 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4\}$ 一定是线性无关的，其中 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 线性无关，且 $\boldsymbol{v}_4$ 不属于 $\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$.

18. 设 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是 $\mathbb{R}^n$ 中的线性无关向量集，证明 $\{\boldsymbol{v}_1,\boldsymbol{v}_1+\boldsymbol{v}_2\}$ 也线性无关.

19. 设 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 是 $\mathbb{R}^3$ 中一条直线上的三个不同点，这条直线不一定经过原点，证明 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 线性相关.

20. 设 $T:\mathbb{R}^n\to\mathbb{R}^m$ 是线性变换，$T(\boldsymbol{u})=\boldsymbol{v}$，证明

$T(-\boldsymbol{u})=-\boldsymbol{v}$.

21. 设 $T:\mathbb{R}^3\to\mathbb{R}^3$ 为线性变换，把每个向量变为它关于平面 $x_2=0$ 的对称点，即 $T(x_1,x_2,x_3)=(x_1,-x_2,x_3)$ ，求出变换 T 的标准矩阵.

22. 设 A 是 3×3 矩阵，线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 把 $\mathbb{R}^3$ 映上到 $\mathbb{R}^3$ ，说明为什么这个变换是一对一的.（提示：考虑主元位置.）

23. 吉温斯旋转是由 $\mathbb{R}^n\to\mathbb{R}^n$ 的一种线性变换，在计算机程序里用来在一个向量中产生一个零元素（通常是矩阵的一列）. $\mathbb{R}^2$ 中的一个吉温斯旋转（见图 1-49）的标准矩阵有形式

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix},\quad a^2+b^2=1$$

求出 a 与 b 把 $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$ 旋转到 $\begin{bmatrix} 5 \\ 0 \end{bmatrix}$.

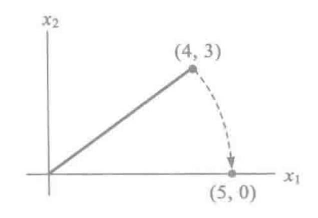


图 1-49　$\mathbb{R}^2$ 中的吉温斯旋转

24. 下列方程是 $\mathbb{R}^3$ 中的吉温斯旋转，求出 a 和 b .

$$\begin{bmatrix} a & 0 & -b \\ 0 & 1 & 0 \\ b & 0 & a \end{bmatrix}\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}=\begin{bmatrix} 2\sqrt{5} \\ 3 \\ 0 \end{bmatrix},\ a^2+b^2=1$$

25. 一幢大的公寓建筑使用模块建筑技术. 每层楼的建筑设计由 3 种设计中选择. A 设计每层有 18 个公寓，包括 3 个三室单元、7 个两室单元和 8 个一室单元；B 设计每层有 4 个三室单元、4 个两室单元和 8 个一室单元；C 设计每层有 5 个三室单元、3 个两室单元和 9 个一室单元. 设该建筑有 x_1 层采取 A 设计，x_2 层采取 B 设计， x_3 层采取 C 设计.

    a. 向量 $x_1\begin{bmatrix} 3 \\ 7 \\ 8 \end{bmatrix}$ 的实际意义是什么？

    b. 写出向量的线性组合表示该建筑所包含的三室、两室和一室单元的总数.

    c. [M]是否可能设计该建筑物，使恰有 66 个三室单元、74 个两室单元和 136 个一室单元？若可能的话，是否有多种方法？说明你的答案.

第2章 矩阵代数

介绍性实例 飞机设计中的计算机模型

为了设计下一代的商业和军用飞机，波音的幻影工作室的工程师们使用三维建模和计算流体动力学. 他们在建造实际的模型之前，研究一个虚拟的模型周围的空气流动，这样做可以很大程度地缩短设计周期，降低成本，而线性代数在这个过程中起了关键的作用.

虚拟的飞机模型的设计从数学的线形轮廓模型开始，它存储在计算机内存中，并可显示在图形显示终端. （图示给出波音 777 的模型.）这个数学模型组织和影响设计和制造飞机外部和内部的每一个过程. 计算流体动力学分析主要考虑的是飞机外部表层的设计.

虽然飞机精巧的外表看上去是光滑的，但其表面的几何曲面是十分复杂的. 除了机翼和机身，飞机上还有引擎机舱、水平尾翼、狭板、襟翼、副翼. 空气在这些结构上的流动决定了飞机在天空中如何运动. 描述气流的方程很复杂，它们必须考虑到引擎的吸气量、引擎的排气量和机翼留下的尾迹. 为了研究气流，工程师们需要飞机表面的精确描述.

用计算机建立飞机外表的模型，首先在原来的线形轮廓模型上添加三维的立方体格子. 这些立方体有的处于飞机的内部，有的在外部，有的和飞机的表面相交. 计算机选出这些相交的立方体，并进一步细分，保留仍然和飞机表面相交的立方体. 这种细分过程一直进行下去，直到立方体非常精细. 一个典型的网格可以含有超过 400 000 个的立方体.

研究飞机表面的气流的过程包含反复求解大型的线性方程组 $A\boldsymbol{x}=\boldsymbol{b}$，涉及的方程和变量个数达到 2 百万个. 向量 $\boldsymbol{b}$ 随来自网格的数据和前面的方程的解而改变. 利用现在商业上买得到的最快的计算机，幻影工作组求解一个气流问题要用数小时至数天的时间. 工作组分析方程组的解之后，会对飞机的外表进行稍微的修改，整个过程又再重新开始，计算流体动力学的分析有可能要进行数千遍.

本章给出协助求解这样大规模方程组的两个重要的概念：

- 分块矩阵：一个典型的计算流体动力学的方程组会有“稀疏”的系数矩阵，上面有许多零元素. 将变量正确地分组会产生有许多零方块的分块矩阵. 2.4 节介绍了这种矩阵及其应用.

- 矩阵分解：即使使用分块矩阵，这样的方程组还是相当的复杂. 为了进一步简化计算，波音的计算流体动力学软件使用了对系数矩阵进行 LU 分解的方法. 2.5 节将讨论 LU 和其他有用的矩阵分解. 关于分解的更详细的内容在本书后面出现.

为了分析气流问题的解，工程师们希望将飞机表面的气流显示出来. 他们利用计算机图形以及线性代数作为图形的引擎. 飞机外表的线形轮廓模型作为许多矩阵数据存储. 图像在计算机屏幕上染色显示后，工程师们可以改变图像的大小，对局部区域进行缩放，以及对图像旋转以看到在视图中被隐藏的部位. 这里的每个操作是通过适当的矩阵乘法运算来实现的. 2.7 节解释了其中的基本思想.图 2-1 为波音公司为机翼做设计.

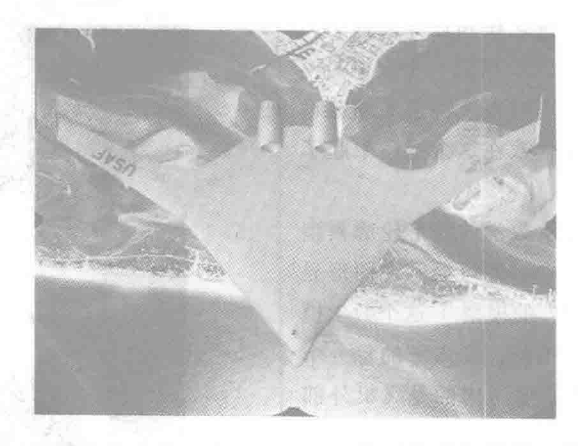


图 2-1 现代计算流体动力学给机翼的设计带来了革新. 波音公司正在为 2020 年或更早的混合机翼和机身的新飞机做设计

当我们学会矩阵的代数运算后，我们分析和解方程的能力将会大大提高. 本章中的定义和定理给出一些基本的工具来处理涉及两个或更多个矩阵的线性代数问题. 对方阵而言，2.3 节的逆矩阵定理把许多以前学过的概念联系在一起. 2.4 节与 2.5 节研究分块矩阵以及矩阵的分解，它们在线性代数的应用很广. 2.6 节和 2.7 节给出了矩阵代数在经济学、计算机图形学中的两个有趣应用.

2.1 矩阵运算

若 A 是 $m\times n$ 矩阵，即有 m 行 n 列的矩阵，A 的第 i 行第 j 列的元素用 a_{ij} 表示，称为 A 的 (i,j) 元素，见图 2-2. 例如，（3,2）元素是在第 3 行第 2 列的数 a_{32}，A 的各列是 $\mathbb{R}^m$ 的向量用（黑体字母）$\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 表示. 当我们特别注意 A 的各列时，我们写成 $A=[\boldsymbol{a}_1\ \ \boldsymbol{a}_2\ \ \cdots\ \ \boldsymbol{a}_n]$，注意 a_{ij} 是第 j 个列向量 $\boldsymbol{a}_j$（从上面算起）的第 i 个元素.

$A=[a_{ij}]$ 的**对角线元素**是 $a_{11},a_{22},a_{33},\cdots$, 它们组成 A 的主对角线. **对角矩阵**是一个方阵，它的非对角线元素全是 0. 元素全是零的 $m\times n$ 矩阵称为**零矩阵**，用 $\mathbf{0}$ 表示. $\mathbf{0}$ 的维数通常可由上下文知道，否则我们就用 $\mathbf{0}_{m\times n}$ 表示.

$$
\begin{array}{r}
\text{第}j\text{列} \\
\text{第}i\text{行}\begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} = A \\
\begin{matrix} \uparrow & \quad & \uparrow & \quad & \uparrow \\ \mathbf{a}_1 & & \mathbf{a}_j & & \mathbf{a}_n \end{matrix}
\end{array}
$$

图 2-2　矩阵记号

和与标量乘法

前面叙述过的向量运算可以自然地推广到矩阵. 我们称两个矩阵相等，若它们有相同的维数（即有相同行数和列数），而且对应元素相等. 若 A 与 B 都是 $m\times n$ 矩阵，则和 $A+B$ 也是 $m\times n$ 矩阵. 它的各列是 A 与 B 对应列之和，因列的向量加法是对应元素相加，$A+B$ 的每个元素也就是 A 与 B 的对应元素相加. 仅当 A 与 B 有相同维数，$A+B$ 才有定义.

例 1　设

$$A=\begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix},\ B=\begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix},\ C=\begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix}$$

则

$$A+B=\begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}$$

但 $A+C$ 没有定义，因 A 与 C 的维数不同. ■

若 r 是标量而 A 是矩阵，则**标量乘法** rA 是一个矩阵，它的每一列是 A 的对应列的 r 倍. 与向量相同，定义 $-A$ 为 $(-1)A$ 而 $A-B$ 为 $A+(-1)B$.

例 2　设 A 与 B 如例 1，则

$$2B=2\begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix}=\begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix}$$

$$A-2B=\begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix}-\begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix}=\begin{bmatrix} 2 & -2 & 3 \\ -7 & -7 & -12 \end{bmatrix}$$ ■

在例 2 中，计算 $A-2B$ 时，不必化为 $A+(-1)2B$，因为通常的代数法则对矩阵的和与标量乘法仍适用，如下列定理所示.

定理 1　设 A, B, C 是相同维数的矩阵，r 与 s 为数，则有

a. $A+B=B+A$　　　　b. $(A+B)+C=A+(B+C)$

c. $A+0=A$　　　　d. $r(A+B)=rA+rB$

e. $(r+s)A=rA+sA$　　　　f. $r(sA)=(rs)A$

为证明定理 1 的各个等式，只要证明左端矩阵和右端矩阵有相同维数且对应各列相等. 维数是无问题的，因 A,B,C 的维数相同. 而由向量的类似性质，立即知道两端的对应各列相等. 例如，若 A,B,C 的第 j 列分别是 $\mathbf{a}_j,\mathbf{b}_j,\mathbf{c}_j$，则 $(A+B)+C$ 与 $A+(B+C)$ 的第 j 列分别是

$$(\boldsymbol{a}_j+\boldsymbol{b}_j)+\boldsymbol{c}_j \text{ 与 } \boldsymbol{a}_j+(\boldsymbol{b}_j+\boldsymbol{c}_j)$$

因对每个 j，这两个向量相等，这就证明了（b）.

由于加法的结合律，我们只要写 $A+B+C$ 即可，它无论按 $(A+B)+C$ 或 $A+(B+C)$ 计算都得同一结果. 同样对四个或更多矩阵也可定义加法.

矩阵乘法

当把矩阵 B 乘以向量 $\boldsymbol{x}$，它将 $\boldsymbol{x}$ 变换为向量 $B\boldsymbol{x}$，若这向量又乘以矩阵 A，结果得向量 $A(B\boldsymbol{x})$，见图 2-3.

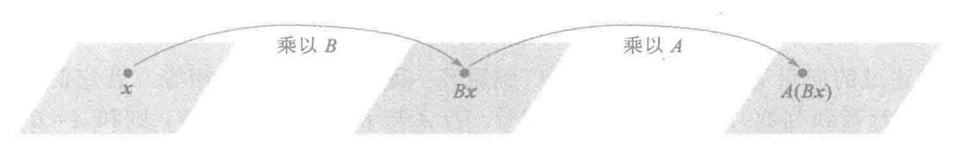


图 2-3　先乘以 B 再乘以 A

于是 $A(B\boldsymbol{x})$ 是由 $\boldsymbol{x}$ 经复合映射变换所得，此映射是 1.7 节所研究的线性变换. 我们的目的是将此复合映射表示为乘以一个矩阵的变换，此矩阵记为 AB，即

$$A(B\boldsymbol{x})=(AB)\boldsymbol{x} \tag{1}$$

见图 2-4.

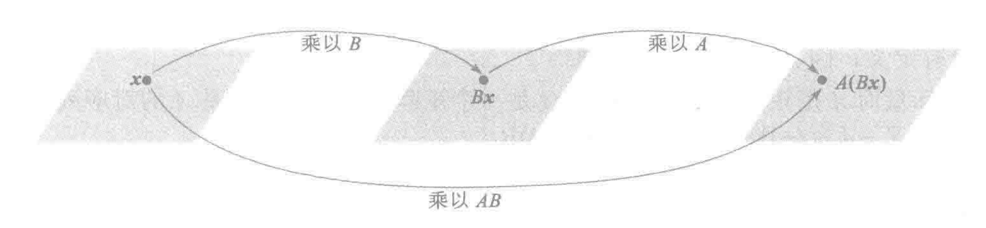


图 2-4　乘以 AB

若 A 是 $m\times n$ 矩阵，B 是 $n\times p$ 矩阵，$\boldsymbol{x}$ 属于 $\mathbb{R}^p$，用 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_p$ 表示 B 的各列，而 $\boldsymbol{x}$ 的元素为 $x_1,\cdots,x_p$，则

$$B\boldsymbol{x}=x_1\boldsymbol{b}_1+\cdots+x_p\boldsymbol{b}_p$$

由于乘以 A 的线性性质

$$A(B\boldsymbol{x})=A(x_1\boldsymbol{b}_1)+\cdots+A(x_p\boldsymbol{b}_p)=x_1A\boldsymbol{b}_1+\cdots+x_pA\boldsymbol{b}_p$$

向量 $A(B\boldsymbol{x})$ 是向量 $A\boldsymbol{b}_1,\cdots,A\boldsymbol{b}_p$ 的线性组合，以 $\boldsymbol{x}$ 的元素为权，若我们把这些向量表示成一个矩阵的各列，就有

$$A(B\boldsymbol{x})=[A\boldsymbol{b}_1\quad A\boldsymbol{b}_2\quad\cdots\quad A\boldsymbol{b}_p]\boldsymbol{x}$$

于是乘以矩阵 $[A\boldsymbol{b}_1\quad A\boldsymbol{b}_2\quad\cdots\quad A\boldsymbol{b}_p]$ 把 $\boldsymbol{x}$ 变为 $A(B\boldsymbol{x})$，我们已经找到了所需要的矩阵.

定义　若 A 是 $m\times n$ 矩阵，B 是 $n\times p$ 矩阵，B 的列是 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_p$，则乘积 AB 是 $m\times p$ 矩阵，它的各列是 $A\boldsymbol{b}_1,\cdots,A\boldsymbol{b}_p$，即

$$AB = A[\boldsymbol{b}_1 \quad \boldsymbol{b}_2 \quad \cdots \quad \boldsymbol{b}_p] = [A\boldsymbol{b}_1 \quad A\boldsymbol{b}_2 \quad \cdots \quad A\boldsymbol{b}_p]$$

这个定义使（1）式对 $\mathbb{R}^p$ 中所有 $\boldsymbol{x}$ 成立，方程（1）证明图 2-4 中的复合映射是线性变换，它的标准矩阵是 AB，矩阵乘法对应线性变换的复合.

例 3 计算 AB，其中 $A=\begin{bmatrix}2 & 3\\ 1 & -5\end{bmatrix}$，$B=\begin{bmatrix}4 & 3 & 6\\ 1 & -2 & 3\end{bmatrix}$.

解 写 $B=[\boldsymbol{b}_1 \ \ \boldsymbol{b}_2 \ \ \boldsymbol{b}_3]$，计算

$$A\boldsymbol{b}_1=\begin{bmatrix}2 & 3\\ 1 & -5\end{bmatrix}\begin{bmatrix}4\\ 1\end{bmatrix}=\begin{bmatrix}11\\ -1\end{bmatrix},\quad A\boldsymbol{b}_2=\begin{bmatrix}2 & 3\\ 1 & -5\end{bmatrix}\begin{bmatrix}3\\ -2\end{bmatrix}=\begin{bmatrix}0\\ 13\end{bmatrix},\quad A\boldsymbol{b}_3=\begin{bmatrix}2 & 3\\ 1 & -5\end{bmatrix}\begin{bmatrix}6\\ 3\end{bmatrix}=\begin{bmatrix}21\\ -9\end{bmatrix}$$

因此

$$AB = A[\boldsymbol{b}_1 \ \ \boldsymbol{b}_2 \ \ \boldsymbol{b}_3]=\begin{bmatrix}11 & 0 & 21\\ -1 & 13 & -9\end{bmatrix}$$

$\uparrow \ \uparrow \ \uparrow$

$A\boldsymbol{b}_1 \ A\boldsymbol{b}_2 \ A\boldsymbol{b}_3$ ■

注意，由 AB 的定义，它的第一列 $A\boldsymbol{b}_1$ 是 A 的各列用 $\boldsymbol{b}_1$ 的各元素为权的线性组合. 其他各列也是这样.

> AB 的每一列都是 A 的各列的线性组合，以 B 的对应列的元素为权.

显然，A 的列数必须等于 B 的行数，才能使线性组合 $A\boldsymbol{b}_1$ 有定义，由定义知，AB 的行数等于 A 的行数，列数等于 B 的列数.

例 4 若 A 是 3×5 矩阵，B 是 5×2 矩阵，AB 和 BA 是否有定义？若有定义，是什么矩阵？

解 因 A 有 5 列，B 有 5 行，乘积 AB 有定义且是 3×2 矩阵：

$$\underset{3\times5}{\overset{A}{\begin{bmatrix}* & * & * & * & *\\ * & * & * & * & *\\ * & * & * & * & *\end{bmatrix}}}\underset{5\times2}{\overset{B}{\begin{bmatrix}* & *\\ * & *\\ * & *\\ * & *\\ * & *\end{bmatrix}}}=\underset{3\times2}{\overset{AB}{\begin{bmatrix}* & *\\ * & *\\ * & *\end{bmatrix}}}$$

相同

AB 的维数

乘积 BA 没有定义，因 B 为 2 列，A 有 3 行. ■

AB 的定义对理论与应用是重要的，但下列法则给出了更有效的计算 AB 的各元素的方法.

> **计算 AB 的行列法则**
>
> 若乘积 AB 有定义，AB 的第 i 行第 j 列的元素是 A 的第 i 行与 B 的第 j 列对应元素乘积之

和. 若 $(AB)_{ij}$ 表示 AB 的 (i,j) 元素，A 为 $m\times n$ 矩阵，则

$$(AB)_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj}$$

为证明这一法则，设 $B=[\boldsymbol{b}_1\ \cdots \boldsymbol{b}_p]$，$AB$ 的第 j 列是 $A\boldsymbol{b}_j$，我们可用 1.4 节的计算 $A\boldsymbol{x}$ 的规则计算 $A\boldsymbol{b}_j$. $A\boldsymbol{b}_j$ 的第 i 个元素是 A 的第 i 行与向量 $\boldsymbol{b}_j$ 的对应元素之积，恰好是上述规则中计算 AB 的 (i,j) 元素的方法.

例 5 使用行列法则计算例 3 中矩阵 AB 的两个元素，观察其中涉及的数会使你更好地理解计算 AB 的两种方法结果相同.

解 要找出 AB 的第 1 行第 3 列的元素，考虑 A 的第 1 行和 B 的第 3 列，把对应元素相乘再加起来，如下所示：

$$AB=\begin{matrix}\rightarrow\\ \\\end{matrix}\begin{bmatrix}2&3\\1&-5\end{bmatrix}\begin{bmatrix}4&3&\overset{\downarrow}{6}\\1&-2&3\end{bmatrix}=\begin{bmatrix}\square&\square&2(6)+3(3)\\\square&\square&\square\end{bmatrix}=\begin{bmatrix}\square&\square&21\\\square&\square&\square\end{bmatrix}$$

对第 2 行第 2 列的元素，用 A 的第 2 行和 B 的第 2 列：

$$\begin{matrix}\\ \rightarrow\end{matrix}\begin{bmatrix}2&3\\1&-5\end{bmatrix}\begin{bmatrix}4&\overset{\downarrow}{3}&6\\1&-2&3\end{bmatrix}=\begin{bmatrix}\square&\square&21\\\square&1(3)+-5(-2)&\square\end{bmatrix}=\begin{bmatrix}\square&\square&21\\\square&13&\square\end{bmatrix}$$

■

例 6 求 AB 的第 2 行，其中

$$A=\begin{bmatrix}2&-5&0\\-1&3&-4\\6&-8&-7\\-3&0&9\end{bmatrix},\quad B=\begin{bmatrix}4&-6\\7&1\\3&2\end{bmatrix}$$

解 由行列法则，AB 的第 2 行是由 A 的第 2 行和 B 的各列相乘所得

$$\begin{matrix}\\ \rightarrow\\ \\ \\ \end{matrix}\begin{bmatrix}2&-5&0\\-1&3&-4\\6&-8&-7\\-3&0&9\end{bmatrix}\begin{bmatrix}\overset{\downarrow}{4}&\overset{\downarrow}{-6}\\7&1\\3&2\end{bmatrix}=\begin{bmatrix}\square&\square\\-4+21-12&6+3-8\\\square&\square\\\square&\square\end{bmatrix}=\begin{bmatrix}\square&\square\\5&1\\\square&\square\\\square&\square\end{bmatrix}$$

■

注意，由例 6 可知，计算 AB 的第 2 行时，我们仅需把 A 的第 2 行写在 B 的左边，得

$$\begin{bmatrix}-1&3&-4\end{bmatrix}\begin{bmatrix}4&-6\\7&1\\3&2\end{bmatrix}=\begin{bmatrix}5&1\end{bmatrix}$$

这在一般情况下也是正确的，可由计算 AB 的行列法则得出. 记 $\text{row}_i(A)$ 表示矩阵 A 的第 i 行，则

$$\text{row}_i(AB)=\text{row}_i(A)\cdot B \tag{1}$$

矩阵乘法的性质

下列定理列出了矩阵乘法的重要性质，I_m 表示 $m\times m$ 单位矩阵，对 $\mathbb{R}^m$ 中的一切 $\boldsymbol{x}$，$I_m\boldsymbol{x}=\boldsymbol{x}$.

定理 2 设 A 为 $m \times n$ 矩阵，B、C 的维数使下列各式的乘积有定义.

a. $A(BC) = (AB)C$ （乘法结合律）

b. $A(B+C) = AB + AC$ （乘法左分配律）

c. $(B+C)A = BA + CA$ （乘法右分配律）

d. $r(AB) = (rA)B = A(rB)$，r 为任意数

e. $I_m A = A = AI_n$ （矩阵乘法的恒等式）

证 性质（b）~（e）在习题中证明，性质（a）是由于矩阵乘法对应于线性变换的复合，而映射的复合是可结合的. 这里给出（a）的另一个基于矩阵乘积的“列”定义的证明. 设

$$C = [\boldsymbol{c}_1 \ \cdots \ \boldsymbol{c}_p]$$

由矩阵乘法的定义

$$BC = [B\boldsymbol{c}_1 \ \cdots \ B\boldsymbol{c}_p]$$

$$A(BC) = [A(B\boldsymbol{c}_1) \ \cdots \ A(B\boldsymbol{c}_p)]$$

由（1）知，AB 的定义使得对一切 $\boldsymbol{x}$ 有 $A(B\boldsymbol{x}) = (AB)\boldsymbol{x}$，所以

$$A(BC) = [(AB)\boldsymbol{c}_1 \ \cdots \ (AB)\boldsymbol{c}_p] = (AB)C$$ ■

定理 1 和定理 2 中的结合律和分配律说明，基本上矩阵表达式中的括号可像实数运算中那样插入和解开，特别地，我们可写乘积 ABC，不管按 $A(BC)$ 或 $(AB)C$[⊖] 计算都相同. 类似地，四个矩阵 $ABCD$ 的乘积可按 $A(BCD)$，$(ABC)D$ 或 $A(BC)D$ 计算，等等. 计算乘积时不管怎样结合都行，但左右顺序必须保持不变.

乘积中的左右顺序是重要的，因为一般来说 AB 与 BA 并不相同. 这并不奇怪，因 AB 的列是 A 的各列的线性组合，而 BA 的各列是 B 的各列的线性组合. 乘积 AB 的因子的位置需要这样强调，即 A 被 B *右乘*，或 B 被 A *左乘*. 若 $AB = BA$，我们称 A 和 B 彼此**可交换**.

例 7 设 $A = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}$ 和 $B = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}$，证明它们不可交换，即证明 $AB \neq BA$.

解

$$AB = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 14 & 3 \\ -2 & -6 \end{bmatrix}$$

$$BA = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 10 & 2 \\ 29 & -2 \end{bmatrix}$$ ■

乘法一般不可交换是矩阵代数与普通实数代数的重要差别，具体例子参阅习题 9~12.

警告

1. 一般情况下，$AB \neq BA$.
2. 消去律对矩阵乘法不成立，即若 $AB = AC$，一般情况下，$B = C$ 并不成立.（见习题 10.）
3. 若乘积 AB 是零矩阵，一般情况下，不能断定 $A = 0$ 或 $B = 0$.（见习题 12.）

⊖ 当 B 是方阵而 C 的列数较 A 的行数少时，计算 $A(BC)$ 比 $(AB)C$ 更方便.

矩阵的乘幂

若 A 是 $n\times n$ 矩阵，k 是正整数，则 A^k 表示 k 个 A 的乘积.

$$A^k = \underbrace{A \cdots A}_{k个}$$

若 A 不是零矩阵，且 $\boldsymbol{x}$ 属于 $\mathbb{R}^n$，则 $A^k\boldsymbol{x}$ 表示 $\boldsymbol{x}$ 被 A 连续左乘 k 次. 若 $k=0$，则 $A^0\boldsymbol{x}$ 就是 $\boldsymbol{x}$ 本身. 因此 A^0 被解释为单位矩阵. 矩阵乘幂在理论和应用中都很有用处（见 2.6 节、4.9 节及本书后面的内容）.

矩阵的转置

给定 $m\times n$ 矩阵 A，则 A 的转置是一个 $n\times m$ 矩阵，用 A^{T} 表示，它的列是由 A 的对应行构成的.

例 8 设

$$A=\begin{bmatrix} a & b \\ c & d \end{bmatrix},\ B=\begin{bmatrix} -5 & 2 \\ 1 & -3 \\ 0 & 4 \end{bmatrix},\ C=\begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & 5 & -2 & 7 \end{bmatrix}$$

则

$$A^{\mathrm{T}}=\begin{bmatrix} a & c \\ b & d \end{bmatrix},\ B^{\mathrm{T}}=\begin{bmatrix} -5 & 1 & 0 \\ 2 & -3 & 4 \end{bmatrix},\ C^{\mathrm{T}}=\begin{bmatrix} 1 & -3 \\ 1 & 5 \\ 1 & -2 \\ 1 & 7 \end{bmatrix}$$

■

定理 3 设 A 与 B 表示矩阵，其维数使下列和与积有定义，则

a. $(A^{\mathrm{T}})^{\mathrm{T}} = A$.

b. $(A+B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}$.

c. 对任意数 r，$(rA)^{\mathrm{T}} = rA^{\mathrm{T}}$.

d. $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$.

（a）～（c）的证明是直接的，这里省略. （d）的证明见习题 33. 通常 $(AB)^{\mathrm{T}}$ 不等于 $A^{\mathrm{T}}B^{\mathrm{T}}$，即使乘积 $A^{\mathrm{T}}B^{\mathrm{T}}$ 是有定义的. 定理 3（d）可推广到多于两个矩阵的乘积，叙述如下：

> 若干个矩阵的乘积的转置等于它们的转置的乘积，但相乘的顺序相反.

习题中包括说明这些性质的例子.

数值计算的注解

1. 在计算机上求出 AB 的最快方法依赖于计算机存储矩阵的方法. 标准的高性能算法（如 LAPACK）中按列计算 AB，正如我们所定义的那样.（一个用 C++语言写成的 LAPACK 版本是按行计算 AB）.

2. AB 的定义使我们可在计算机上用并行算法计算，B 的列可单独或分组分配给不同的处理器，因此可以同时计算 AB 的各列.

练习题

1. 因 $\mathbb{R}^n$ 中向量可以看作 $n\times1$ 矩阵，转置矩阵的性质也适用于向量，令

$$A=\begin{bmatrix}1 & -3\\ -2 & 4\end{bmatrix} \text{和}\ \boldsymbol{x}=\begin{bmatrix}5\\ 3\end{bmatrix}$$

计算 $(A\boldsymbol{x})^{\mathrm{T}},\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}},\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}$ 和 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$． $A^{\mathrm{T}}\boldsymbol{x}^{\mathrm{T}}$ 是否有定义？

2. 设 A 为 4×4 向量， $\boldsymbol{x}$ 是 $\mathbb{R}^4$ 中向量．计算 $A^2\boldsymbol{x}$ 的最快方法是什么？计算乘法的次数．

习题 2.1

习题 1~2 中，计算矩阵的和或乘积，如果它没有定义，则说明理由．设

$$A=\begin{bmatrix}2 & 0 & -1\\ 4 & -5 & 2\end{bmatrix},B=\begin{bmatrix}7 & -5 & 1\\ 1 & -4 & -3\end{bmatrix},$$

$$C=\begin{bmatrix}1 & 2\\ -2 & 1\end{bmatrix},D=\begin{bmatrix}3 & 5\\ -1 & 4\end{bmatrix},E=\begin{bmatrix}-5\\ 3\end{bmatrix}$$

1. $-2A,\ B-2A,\ AC,\ CD$
2. $A+3B,\ 2C-3E,\ DB,\ EC$

下面的习题中，假设每个矩阵表达式是有定义的，矩阵（和向量）的维数是相互匹配的．

3. 设 $A=\begin{bmatrix}2 & -5\\ 3 & -2\end{bmatrix}$，计算 $3I_2-A$ 及 $(3I_2)A$．
4. 计算 $A-5I_3$ 及 $(5I_3)A$，其中

$$A=\begin{bmatrix}5 & -1 & 3\\ -4 & 3 & -6\\ -3 & 1 & 2\end{bmatrix}$$

习题 5~6 中，用两种方法计算乘积 AB；(a) 根据定义分别计算 $A\boldsymbol{b}_1$ 及 $A\boldsymbol{b}_2$，(b) 利用计算 AB 的行列法则．

5. $A=\begin{bmatrix}-1 & 3\\ 2 & 4\\ 5 & -3\end{bmatrix},\ B=\begin{bmatrix}4 & -2\\ -2 & 3\end{bmatrix}$
6. $A=\begin{bmatrix}4 & -3\\ -3 & 5\\ 0 & 1\end{bmatrix},\ B=\begin{bmatrix}1 & 4\\ 2 & -2\end{bmatrix}$
7. 若矩阵 A 是 5×3，乘积 AB 是 5×7， B 的维数是多少？
8. 若 BC 是 5×4 矩阵， B 有几行？
9. 设 $A=\begin{bmatrix}2 & 3\\ -1 & 1\end{bmatrix}$ 和 $B=\begin{bmatrix}1 & 9\\ -3 & k\end{bmatrix}$， k 取什么值时 $AB=BA$？
10. 设 $A=\begin{bmatrix}3 & -6\\ -1 & 2\end{bmatrix},B=\begin{bmatrix}-1 & 1\\ 3 & 4\end{bmatrix},C=\begin{bmatrix}-3 & -5\\ 2 & 1\end{bmatrix}$，证明 $AB=AC$ 但 $B\neq C$．
11. 设 $A=\begin{bmatrix}1 & 2 & 3\\ 2 & 4 & 5\\ 3 & 5 & 6\end{bmatrix},D=\begin{bmatrix}5 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & 2\end{bmatrix}$，计算 AD 和 DA，说明当 A 右乘或左乘以 D 时， A 的行或列如何变化，求出 3×3 对角矩阵 B，不是单位矩阵或零矩阵，使 $AB=BA$．
12. 设 $A=\begin{bmatrix}3 & -6\\ -2 & 4\end{bmatrix}$，求 2×2 矩阵 B 使 $AB=\boldsymbol{0}$，要求 B 有两个不相同的非零列．
13. 设 $\boldsymbol{r}_1,\cdots,\boldsymbol{r}_p$ 为 $\mathbb{R}^n$ 中向量，Q 为 $m\times n$ 矩阵，把矩阵 $[Q\boldsymbol{r}_1\ \ \cdots\ \ Q\boldsymbol{r}_p]$ 写成两个矩阵的乘积(任何一个矩阵都不是单位矩阵）．
14. 设 U 是 1.8 节例 6 所描述的 3×2 成本矩阵．U 的第一列给出生产产品 B 每美元产出的成本，而第二列给出生产产品 C 每美元产出的成本．(成本分为材料、劳动和管理.）设 $\boldsymbol{q}_1$ 是 $\mathbb{R}^2$ 中向量，给出一年第一季度生产产品 B 和 C（以美元计算）的产出， $\boldsymbol{q}_2,\boldsymbol{q}_3,\boldsymbol{q}_4$ 是类似的向量，分别给出一年第二、三和四季度生产产品 B 和 C 的产出.给出矩阵 UQ 中的数据的经济解释，其中 $Q=[\boldsymbol{q}_1\ \ \boldsymbol{q}_2\ \ \boldsymbol{q}_3\ \ \boldsymbol{q}_4]$．

习题 15~16 中的矩阵 A,B,C 使所说的加法和乘法运算能够进行．标出每个命题的真假，给出理由．

15. a. 若 A,B 为 2×2 矩阵，它们的列分别为 $\boldsymbol{a}_1,\boldsymbol{a}_2$ 和 $\boldsymbol{b}_1,\boldsymbol{b}_2$，则 $AB=[\boldsymbol{a}_1\boldsymbol{b}_1\ \ \boldsymbol{a}_2\boldsymbol{b}_2]$．

b. AB 的每一列是 B 的列的线性组合，并以 A 的对应列作为权.

c. $AB+AC=A(B+C)$.

d. $A^{\mathrm{T}}+B^{\mathrm{T}}=(A+B)^{\mathrm{T}}$.

e. 矩阵的乘积的转置等于它们的转置的乘积.

16. a. AB 的第 1 行是 A 的第 1 行被 B 右乘.

b. 若 A 与 B 为 3×3 矩阵，$B=[\boldsymbol{b}_1 \quad \boldsymbol{b}_2 \quad \boldsymbol{b}_3]$，则 $AB=[A\boldsymbol{b}_1+A\boldsymbol{b}_2+A\boldsymbol{b}_3]$.

c. 若 A 是 $n\times n$ 矩阵，则 $(A^2)^{\mathrm{T}}=(A^{\mathrm{T}})^2$.

d. $(ABC)^{\mathrm{T}}=C^{\mathrm{T}}A^{\mathrm{T}}B^{\mathrm{T}}$.

e. 矩阵和的转置等于它们的转置的和.

17. 若 $A=\begin{bmatrix}1 & -3\\ -3 & 5\end{bmatrix}$ 和 $AB=\begin{bmatrix}-3 & -11\\ 1 & 17\end{bmatrix}$，确定 B 的第一列与第二列.

18. 设 B 的第 3 列全是零，那么 AB 的第 3 列如何？

19. 设 B 的第 3 列是前 2 列的和，那么 AB 的第 3 列如何？为什么？

20. 设 B 的前两列 $\boldsymbol{b}_1$ 和 $\boldsymbol{b}_2$ 相等. 那么 AB 的各列如何（设 AB 有定义）？为什么？

21. 设 AB 的最后一列全是零，但 B 本身没有零列，那么 A 的各列如何？

22. 若 B 的各列线性相关，证明 AB 的各列也线性相关.

23. 设 $CA=I_n$（$n\times n$ 单位矩阵），证明方程 $A\boldsymbol{x}=\boldsymbol{0}$ 只有平凡解.解释为什么 A 的列数不可以多于行数.

24. 设 A 是 $3\times n$ 矩阵，其列生成 $\mathbb{R}^3$，解释如何构造一个 $n\times 3$ 矩阵 D，使得 $AD=I_3$.

25. 设 A 是 $m\times n$ 矩阵，存在 $n\times m$ 矩阵 C 和 D，使 $CA=I_n$ 和 $AD=I_m$. 证明 $m=n$ 和 $C=D$.（提示：利用乘积 CAD.）

26. 设 $AD=I_m$（$m\times m$ 单位矩阵），证明，对任意 $\mathbb{R}^m$ 中的 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 有解.（提示：利用方程 $AD\boldsymbol{b}=\boldsymbol{b}$.）解释为什么 A 的行数不可以多于列数.

27 题和 28 题中，把 $\mathbb{R}^n$ 中向量看作 $n\times 1$ 矩阵，对 $\mathbb{R}^n$ 中 $\boldsymbol{u},\boldsymbol{v}$，矩阵乘积 $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}$ 是 1×1 矩阵，称为 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的**数量积或内积**，它通常当作实数而省略括号. 矩阵乘积 $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ 是 $n\times n$ 矩阵，称为 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的**外积**. 这些乘积 $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}$ 和 $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ 以后将用到.

27. 设 $\boldsymbol{u}=\begin{bmatrix}-3\\ 2\\ -5\end{bmatrix},\boldsymbol{v}=\begin{bmatrix}a\\ b\\ c\end{bmatrix}$，计算 $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v},\boldsymbol{v}^{\mathrm{T}}\boldsymbol{u},\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ 和 $\boldsymbol{v}\boldsymbol{u}^{\mathrm{T}}$.

28. 若 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 属于 $\mathbb{R}^n$，$\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}$ 和 $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{u}$ 有什么关系？$\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ 和 $\boldsymbol{v}\boldsymbol{u}^{\mathrm{T}}$ 有什么关系？

29. 证明定理 2（b）和 2（c），应用行列法则，$A(B+C)$ 的 (i,j) 元素可写成

$$a_{i1}(b_{1j}+c_{1j})+\cdots+a_{in}(b_{nj}+c_{nj}) \quad 或$$

$$\sum_{k=1}^{n}a_{ik}(b_{kj}+c_{kj})$$

30. 证明定理 2（d）.（提示：$(rA)B$ 的 (i,j) 元素是 $(ra_{i1})b_{1j}+\cdots+(ra_{in})b_{nj}$.）

31. 证明 $I_mA=A,A$ 为 $m\times n$ 矩阵. 假设对所有 $\mathbb{R}^m$ 中的 $\boldsymbol{x}$，有 $I_m\boldsymbol{x}=\boldsymbol{x}$.

32. 证明 $AI_n=A,A$ 为 $m\times n$ 矩阵.（提示：应用 AI_n 的列定义）.

33. 证明定理 3（d）.（提示：考虑 $(AB)^{\mathrm{T}}$ 的第 j 行.）

34. 给出 $(AB\boldsymbol{x})^{\mathrm{T}}$ 的公式，其中 $\boldsymbol{x}$ 是向量，A 和 B 是适当维数的矩阵.

35. [M]读矩阵程序的文件，写出产生下列矩阵的命令（不要键入矩阵的任何一个元素）.

a. 4×5 的零矩阵.

b. 5×3 的矩阵，其元素都是 1.

c. 5×5 单位矩阵.

d. 4×4 对角矩阵，对角元素是 3，4，2，5.

检验矩阵代数的新思想或做猜想的一个有用的方法，是使用随机选择的矩阵进行计算. 用一些矩阵验证一个性质并不能证明在一般情况下这个性质是成立的，但这样做有助于对性质的理解，如果该性质是假的，你会在一些计算之后发现.

36. [M]写出生成 5×6 矩阵，其元素为随机数的命令，这些随机数在什么范围内？说出如何生成随机 4×4 矩阵，它的元素为整数且在−9 和 9

之间.（若 x 是随机数，$0<x<1$，则 $-9.5<19(x-0.5)<9.5$.）

37. [M]构造 4×4 随机矩阵 A 和 B，检验 $AB=BA$ 是否成立. 最好的方法是计算 $AB-BA$ 并证明它等于零矩阵. 对三对随机 4×4 矩阵检验 $AB-BA$ 是否成立.

38. [M]构造一个 5×5 随机矩阵 A，检验 $(A+I)\cdot(A-I)=A^2-I$ 是否成立. 最好的办法是计算 $(A+I)(A-I)-(A^2-I)$ 并证明它等于零矩阵. 对三个随机矩阵进行检验. 对三对随机 4×4 矩阵检验 $(A+B)(A-B)=A^2-B^2$ 是否成立.

39. [M]用至少三对 4×4 矩阵 A 和 B 检验等式 $(A+B)^{\mathrm{T}}=A^{\mathrm{T}}+B^{\mathrm{T}}$ 和 $(AB)^{\mathrm{T}}=A^{\mathrm{T}}B^{\mathrm{T}}$ 是否成立.（见习题 37.）（注：多数矩阵程序用 A' 表示 A^{T}.）

40. [M]设

$$S=\begin{bmatrix}0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\\0&0&0&0&0\end{bmatrix}$$

对 $k=2,\cdots,6$，计算 S^k.

41. [M]叙述当你计算 A^5, A^{10}, A^{20} 和 A^{30} 时发生的情况.

$$A=\begin{bmatrix}1/4&1/2&1/4\\1/2&1/3&1/6\\1/4&1/6&7/12\end{bmatrix}$$

练习题答案

1. $Ax=\begin{bmatrix}1&-3\\-2&4\end{bmatrix}\begin{bmatrix}5\\3\end{bmatrix}=\begin{bmatrix}-4\\2\end{bmatrix}$，所以 $(A\boldsymbol{x})^{\mathrm{T}}=[-4\quad 2]$，同样

$$\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}=[5\ \ 3]\begin{bmatrix}1&-2\\-3&4\end{bmatrix}=[-4\ \ 2]$$

如定理 3（d）所说，$(A\boldsymbol{x})^{\mathrm{T}}$ 和 $\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}$ 相等. 其次

$$\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}=\begin{bmatrix}5\\3\end{bmatrix}[5\quad 3]=\begin{bmatrix}25&15\\15&9\end{bmatrix}$$

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=[5\quad 3]\begin{bmatrix}5\\3\end{bmatrix}=[25+9]=34$$

如 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$ 的 1×1 矩阵通常不写括号. 最后 $A^{\mathrm{T}}\boldsymbol{x}^{\mathrm{T}}$ 没有定义，因 $\boldsymbol{x}^{\mathrm{T}}$ 没有 2 行与 A^{T} 的两列配合.

2. 最快的方法是按 $A(A\boldsymbol{x})$ 计算 $A^2\boldsymbol{x}$. 计算 $A\boldsymbol{x}$ 需要 16 次乘法，每个元素需要 4 次. 计算 $A(A\boldsymbol{x})$ 还需 16 次. 计算 A^2 需要 64 次乘法，因 A^2 有 16 个元素，每个需 4 次. 之后 $A^2\boldsymbol{x}$ 还需 16 次乘法，总共需 80 次.

2.2 矩阵的逆

矩阵代数提供了对矩阵方程进行运算的工具以及许多与普通的实数代数相似的有用公式. 本节研究矩阵中与非零数的倒数（即乘法逆）类似的问题.

实数 5 的乘法逆是 1/5 或 5^{-1}，它满足方程

$$5^{-1}\cdot 5=1 \text{ 和 } 5\cdot 5^{-1}=1$$

矩阵对逆的一般化也要求两个方程同时成立，并避免使用斜线记号表示除法，因为矩阵乘法不是可交换的. 进一步，完全的一般化是可能的，当且仅当有关矩阵是方阵.[⊖]

⊖ 也许有人会说一个 $m\times n$ 矩阵 A 是可逆的，如果存在 $n\times m$ 矩阵 C 和 D 使 $CA=I_n$ 且 $AD=I_m$. 但是，这两个方程可推出 A 是方阵，且 $C=D$. 因此 A 是可逆的定义同上. 见 2.1 节习题 23~25.

一个 $n\times n$ 矩阵 A 是**可逆的**，若存在一个 $n\times n$ 矩阵 C 使

$$AC=I \quad 且 \quad CA=I$$

这里 $I=I_n$ 是 $n\times n$ 单位矩阵. 这时称 C 是 A 的**逆阵**. 实际上，C 由 A 唯一确定，因为若 B 是另一个 A 的逆阵，那么将有 $B=BI=B(AC)=(BA)C=IC=C$，于是，若 A 可逆，它的逆是唯一的，我们将它记为 A^{-1}，于是

$$AA^{-1}=I \quad 且 \quad A^{-1}A=I$$

不可逆矩阵有时称为**奇异矩阵**，而可逆矩阵也称为**非奇异矩阵**.

例 1 若 $A=\begin{bmatrix}2 & 5\\ -3 & -7\end{bmatrix}$, $C=\begin{bmatrix}-7 & -5\\ 3 & 2\end{bmatrix}$，则

$$AC=\begin{bmatrix}2 & 5\\ -3 & -7\end{bmatrix}\begin{bmatrix}-7 & -5\\ 3 & 2\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

$$CA=\begin{bmatrix}-7 & -5\\ 3 & 2\end{bmatrix}\begin{bmatrix}2 & 5\\ -3 & -7\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

所以 $C=A^{-1}$. ■

这里给出 2×2 矩阵可逆的检验方法，同时给出一个简单的公式给出它的逆矩阵.

定理 4 *设 $A=\begin{bmatrix}a & b\\ c & d\end{bmatrix}$，若 $ad-bc\neq 0$，则 A 可逆且*

$$A^{-1}=\frac{1}{ad-bc}\begin{bmatrix}d & -b\\ -c & a\end{bmatrix}$$

若 $ad-bc=0$，则 A 不可逆.

定理 4 的简单证明见 25 题与 26 题. 数 $ad-bc$ 称为 A 的行列式，记为

$$\det A=ad-bc$$

定理 4 说明，2×2 矩阵 A 可逆，当且仅当 $\det A\neq 0$.

例 2 求 $A=\begin{bmatrix}3 & 4\\ 5 & 6\end{bmatrix}$ 的逆.

解 因为 $\det A=3(6)-4(5)=-2\neq 0$，所以 A 可逆且

$$A^{-1}=\frac{1}{-2}\begin{bmatrix}6 & -4\\ -5 & 3\end{bmatrix}=\begin{bmatrix}6/(-2) & -4/(-2)\\ -5/(-2) & 3/(-2)\end{bmatrix}=\begin{bmatrix}-3 & 2\\ 5/2 & -3/2\end{bmatrix}$$ ■

可逆矩阵在线性代数中是很重要的——主要用在计算和公式推导中，如下定理所示，有时逆矩阵在实际应用中也会出现，如下面例 3 所示.

定理 5 *若 A 是可逆 $n\times n$ 矩阵，则对每一 $\mathbb{R}^n$ 中的 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 有唯一解 $\boldsymbol{x}=A^{-1}\boldsymbol{b}$.*

证 取 $\mathbb{R}^n$ 中任意一个 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 有解，因若以 $A^{-1}\boldsymbol{b}$ 代 $\boldsymbol{x}$，有 $A\boldsymbol{x}=A(A^{-1}\boldsymbol{b})=(AA^{-1})\boldsymbol{b}=I\boldsymbol{b}=\boldsymbol{b}$，所以 $A^{-1}\boldsymbol{b}$ 是解. 为证明解是唯一的，我们证明若 $\boldsymbol{u}$ 是一个解，则 $\boldsymbol{u}$ 必是 $A^{-1}\boldsymbol{b}$，事实上，若 $A\boldsymbol{u}=\boldsymbol{b}$，两边同乘 A^{-1} 得

$$A^{-1}Au = A^{-1}\boldsymbol{b},\ I\boldsymbol{u} = A^{-1}\boldsymbol{b},\ \boldsymbol{u} = A^{-1}\boldsymbol{b}$$ ■

例 3 一条水平的弹性梁的两端支柱，在点 1，2，3 受力作用，如图 2-5 所示，设 $\mathbb{R}^3$ 中的 $\boldsymbol{f}$ 表示它在这三点受的力，$\boldsymbol{y}$ 为梁在这三点的形变. 利用虎克定律，可以证明

$$\boldsymbol{y} = D\boldsymbol{f}$$

这里 D 称为弹性矩阵. 它的逆称为刚性矩阵，说明 D 与 D^{-1} 各列的物理意义.

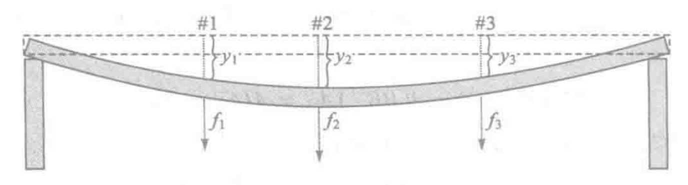


图 2-5 弹性梁的形变

解 记 $I_3 = [\boldsymbol{e}_1\ \ \boldsymbol{e}_2\ \ \boldsymbol{e}_3]$，有 $D = DI_3 = [D\boldsymbol{e}_1\ \ D\boldsymbol{e}_2\ \ D\boldsymbol{e}_3]$，向量 $\boldsymbol{e}_1 = (1,0,0)$ 表示 1 单位力向下作用于点 1（其他两点的力为零），$D\boldsymbol{e}_1$（即 D 的第一列）表示在点 1 处施加 1 单位力产生的梁的形变. 类似地，可以说明 D 的第二和第三列.

为研究刚性矩阵，注意到方程 $\boldsymbol{f} = D^{-1}\boldsymbol{y}$ 计算形变向量 $\boldsymbol{y}$ 给定时的力向量 $\boldsymbol{f}$. 记 $D^{-1} = D^{-1}I_3 = \left[D^{-1}\boldsymbol{e}_1\ \ D^{-1}\boldsymbol{e}_2\ \ D^{-1}\boldsymbol{e}_3\right]$，现解释 $\boldsymbol{e}_1$ 为形变向量，于是 $D^{-1}\boldsymbol{e}_1$ 给出产生这个形变的力. 因此，D^{-1} 的第一列表示，为了使点 1 的形变为 1 单位，其他两点形变为 0，所需要作用的力. 类似地，D^{-1} 的第二和第三列分别表示为了在点 2 和 3 产生 1 单位的形变所需要作用的力. 在每一列中，其中一点或两点作用的力必须为负值（指向上），以在指定的点产生单位形变而其他点没有形变. 若弹性用每磅力产生的形变的英寸数衡量，则刚性矩阵的元素是每英寸形变所需力的磅数. ■

定理 5 的公式很少用来解方程 $A\boldsymbol{x} = \boldsymbol{b}$，因为 $[A\ \ \boldsymbol{b}]$ 的行变换通常更快（当计算有舍入误差时，行变换也更精确.）一个可能的例外是 2×2 矩阵. 这时用 A^{-1} 的公式进行心算会更容易，如下例所示.

例 4 用例 2 中矩阵 A 的逆矩阵解方程组

$$3x_1 + 4x_2 = 3$$
$$5x_1 + 6x_2 = 7$$

解 该方程组就是 $A\boldsymbol{x} = \boldsymbol{b}$，所以

$$\boldsymbol{x} = A^{-1}\boldsymbol{b} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix}\begin{bmatrix} 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$$ ■

下列定理给出可逆矩阵的三个有用事实.

定理 6

a. *若 A 是可逆矩阵，则 A^{-1} 也可逆而且 $(A^{-1})^{-1} = A$.*

b. *若 A 和 B 都是 $n \times n$ 可逆矩阵，AB 也可逆，且其逆是 A 和 B 的逆矩阵按相反顺序的乘积，*

即

$$(AB)^{-1}=B^{-1}A^{-1}$$

c. 若 A 可逆，则 A^{T} 也可逆，且其逆是 A^{-1} 的转置，即 $(A^{\mathrm{T}})^{-1}=(A^{-1})^{\mathrm{T}}$.

证 为证明（a），我们需要找矩阵 C 使

$$A^{-1}C=I \quad 且 \quad CA^{-1}=I$$

显然 A 满足这些方程. 因此 A^{-1} 可逆且 A 是它的逆阵. 下一步，为证明（b），我们应用乘法结合律：

$$(AB)(B^{-1}A^{-1})=A(BB^{-1})A^{-1}=AIA^{-1}=AA^{-1}=I$$

类似地，可以证明 $(B^{-1}A^{-1})(AB)=I$，因此 AB 是可逆的，且其逆为 $B^{-1}A^{-1}$. 对于（c），利用定理3（d），公式从右向左，有 $(A^{-1})^{\mathrm{T}}A^{\mathrm{T}}=(AA^{-1})^{\mathrm{T}}=I^{\mathrm{T}}=I$. 类似地，$A^{\mathrm{T}}(A^{-1})^{\mathrm{T}}=I^{\mathrm{T}}=I$. 因此 A^{T} 是可逆的，其逆是 $(A^{-1})^{\mathrm{T}}$. ■

定理6（b）的下列推广以后要用到.

> 若干个 $n\times n$ 可逆矩阵的积也是可逆的，其逆等于这些矩阵的逆按相反顺序的乘积.

在可逆矩阵与矩阵的行变换之间有一种重要的联系，它引出了计算逆矩阵的一种方法，我们将看到，可逆矩阵行等价于单位矩阵，而我们可通过观察 A 行化简为 I 这一过程求出 A^{-1}.

初等矩阵

把单位矩阵进行一次行变换，就得到**初等矩阵**. 下列例子说明三种初等矩阵.

例5 设

$$E_1=\begin{bmatrix}1&0&0\\0&1&0\\-4&0&1\end{bmatrix},\ E_2=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix},\ E_3=\begin{bmatrix}1&0&0\\0&1&0\\0&0&5\end{bmatrix},A=\begin{bmatrix}a&b&c\\d&e&f\\g&h&i\end{bmatrix}$$

计算 E_1A, E_2A 与 E_3A，说明这些乘积可由 A 进行变换得到.

解 我们有

$$E_1A=\begin{bmatrix}a&b&c\\d&e&f\\g-4a&h-4b&i-4c\end{bmatrix},\ E_2A=\begin{bmatrix}d&e&f\\a&b&c\\g&h&i\end{bmatrix},\ E_3A=\begin{bmatrix}a&b&c\\d&e&f\\5g&5h&5i\end{bmatrix}$$

把 A 的第1行乘–4加到第3行得 E_1A（这是倍加行变换），交换 A 的第1行与第2行得 E_2A，把 A 的第3行乘以5得 E_3A. ■

把 $3\times n$ 矩阵左乘以（即在左边相乘）例5中的 E_1 也有相同的结果，即把第1行的–4倍加到第3行. 特别地，$E_1I=E_1$，我们看到，E_1 本身是把单位矩阵以同一行变换作用所得. 于是例5说明了下列关于初等矩阵的一般事实，见习题27和28.

> 若对 $m\times n$ 矩阵 A 进行某种初等行变换，所得矩阵可写成 EA，其中 E 是 $m\times m$ 矩阵，是由 I_m 进行同一行变换所得.

因为行变换是可逆的，如我们在 1.1 节所示，初等矩阵也是可逆的. 若 E 是由 I 进行变换所得，则有同一类型的另一行变换把 E 变回 I. 因此，有初等矩阵 F 使 $FE=I$. 因 E 和 F 对应互逆的变换，所以也有 $EF=I$.

每个初等矩阵 E 是可逆的，E 的逆是一个同类型的初等矩阵，它把 E 变回 I.

例 6 求 $E_1=\begin{bmatrix}1&0&0\\0&1&0\\-4&0&1\end{bmatrix}$ 的逆.

解 为把 E_1 变成 I，把第 1 行的 4 倍加上第 3 行，这相应于初等矩阵

$$E_1^{-1}=\begin{bmatrix}1&0&0\\0&1&0\\+4&0&1\end{bmatrix}$$ ■

下列定理给出了判断矩阵可逆的方法，也给出计算逆矩阵的方法.

定理 7 $n\times n$ 矩阵 A 是可逆的，当且仅当 A 行等价于 I_n，这时，把 A 变为 I_n 的一系列初等行变换同时把 I_n 变成 A^{-1}.

证 设 A 是可逆矩阵，则对任意 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 有解（定理 5），A 在每一行有主元位置（1.4 节定理 4），因 A 是方阵，这 n 个主元位置必在对角线上. 这就是说 A 的简化阶梯形是 I_n，即 $A\sim I_n$.

反之，若 $A\sim I_n$，因每一步行变换对应于左乘一个初等矩阵，就是说，存在初等矩阵 $E_1,\cdots,E_p$ 使

$$A\sim E_1A\sim E_2(E_1A)\sim\cdots\sim E_p(E_{p-1}\cdots E_1A)=I_n$$

即

$$E_pE_{p-1}\cdots E_1A=I_n \tag{1}$$

因为 $E_p\cdots E_1$ 是可逆矩阵的乘积，因此也是可逆矩阵，由（1）式推出

$$(E_p\cdots E_1)^{-1}(E_p\cdots E_1)A=(E_p\cdots E_1)^{-1}I_n$$
$$A=(E_p\cdots E_1)^{-1}$$

于是 A 是可逆的，因它是可逆矩阵的逆（定理 6），同样有

$$A^{-1}=[(E_p\cdots E_1)^{-1}]^{-1}=E_p\cdots E_1$$

于是 $A^{-1}=E_p\cdots E_1\cdot I_n$，这就是说，$A^{-1}$ 可由依次以 $E_1,\cdots,E_p$ 作用于 I_n 而得到，它们就是（1）式中把 A 变为 I_n 的同一行变换序列. ■

求 A^{-1} 的算法

若我们把 A 和 I 排在一起构成增广矩阵 $[A\ \ I]$，则对此矩阵进行行变换时，A 和 I 受到同一变换，由定理 7，要么有一系列的行变换把 A 变成 I，同时把 I 变成 A^{-1}，要么 A 是不可逆的.

求 A^{-1} 的算法

把增广矩阵[A I]进行行化简. 若 A 行等价于 I，则[A I]行等价于[I A^{-1}]，否则 A 没有逆.

例 7 求矩阵 $A=\begin{bmatrix}0&1&2\\1&0&3\\4&-3&8\end{bmatrix}$ 的逆，假如它存在的话.

解

$$[A\ \ I]=\begin{bmatrix}0&1&2&1&0&0\\1&0&3&0&1&0\\4&-3&8&0&0&1\end{bmatrix}\sim\begin{bmatrix}1&0&3&0&1&0\\0&1&2&1&0&0\\4&-3&8&0&0&1\end{bmatrix}$$

$$\sim\begin{bmatrix}1&0&3&0&1&0\\0&1&2&1&0&0\\0&-3&-4&0&-4&1\end{bmatrix}\sim\begin{bmatrix}1&0&3&0&1&0\\0&1&2&1&0&0\\0&0&2&3&-4&1\end{bmatrix}$$

$$\sim\begin{bmatrix}1&0&3&0&1&0\\0&1&2&1&0&0\\0&0&1&3/2&-2&1/2\end{bmatrix}\sim\begin{bmatrix}1&0&0&-9/2&7&-3/2\\0&1&0&-2&4&-1\\0&0&1&3/2&-2&1/2\end{bmatrix}$$

因为 $A\sim I$，由定理 7 知 A 可逆，且

$$A^{-1}=\begin{bmatrix}-9/2&7&-3/2\\-2&4&-1\\3/2&-2&1/2\end{bmatrix}$$

最好把答案验证一下：

$$AA^{-1}=\begin{bmatrix}0&1&2\\1&0&3\\4&-3&8\end{bmatrix}\begin{bmatrix}-9/2&7&-3/2\\-2&4&-1\\3/2&-2&1/2\end{bmatrix}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$$

因 A 可逆，不必验证 $A^{-1}A=I$. ■

逆矩阵的另一个观点

用 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n$ 表示 I_n 的各列. 则把[A I]行变换成[I A^{-1}]的过程可看作解 n 个方程组.

$$A\boldsymbol{x}=\boldsymbol{e}_1,\ A\boldsymbol{x}=\boldsymbol{e}_2,\ \cdots,\ A\boldsymbol{x}=\boldsymbol{e}_n \tag{2}$$

其中这些方程组的“增广列”都放在 A 的右边，构成矩阵

$$[A\ \ \boldsymbol{e}_1\ \ \boldsymbol{e}_2\ \ \cdots\ \ \boldsymbol{e}_n]=[A\ \ I]$$

方程 $AA^{-1}=I$ 及矩阵乘法的定义说明 A^{-1} 的列正好是方程（2）的解. 这一点是很有用的，因为在某些应用问题中，只需要 A^{-1} 的一列或两列. 这时只需要解（2）中的相应方程.

数值计算的注解 在实际中，很少计算 A^{-1}，除非需要 A^{-1} 的元素. 计算 A^{-1} 和 $A^{-1}\boldsymbol{b}$ 总共需要的运算次数大约是用行变换解方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的 3 倍，而且行变换可能更为精确.

练习题

1. 应用行列式判断以下矩阵是否可逆：

a. $\begin{bmatrix} 3 & -9 \\ 2 & 6 \end{bmatrix}$ b. $\begin{bmatrix} 4 & -9 \\ 0 & 5 \end{bmatrix}$ c. $\begin{bmatrix} 6 & -9 \\ -4 & 6 \end{bmatrix}$

2. 求 $A=\begin{bmatrix} 1 & -2 & -1 \\ -1 & 5 & 6 \\ 5 & -4 & 5 \end{bmatrix}$ 的逆矩阵，假如它存在.

习题 2.2

求习题 1~4 中矩阵的逆.

1. $\begin{bmatrix} 8 & 6 \\ 5 & 4 \end{bmatrix}$

2. $\begin{bmatrix} 3 & 2 \\ 8 & 5 \end{bmatrix}$

3. $\begin{bmatrix} 7 & 3 \\ -6 & -3 \end{bmatrix}$

4. $\begin{bmatrix} 2 & -4 \\ 4 & -6 \end{bmatrix}$

5. 用习题 1 求出的逆矩阵解下列方程组

$$\begin{aligned} 8x_1 + 6x_2 &= 2 \\ 5x_1 + 4x_2 &= -1 \end{aligned}$$

6. 用习题 3 求出的逆矩阵解下列方程组

$$\begin{aligned} 7x_1 + 3x_2 &= -9 \\ -6x_1 - 3x_2 &= 4 \end{aligned}$$

7. 设 $A=\begin{bmatrix} 1 & 2 \\ 5 & 12 \end{bmatrix}$, $\boldsymbol{b}_1=\begin{bmatrix} -1 \\ 3 \end{bmatrix}$, $\boldsymbol{b}_2=\begin{bmatrix} 1 \\ -5 \end{bmatrix}$, $\boldsymbol{b}_3=\begin{bmatrix} 2 \\ 6 \end{bmatrix}$, $\boldsymbol{b}_4=\begin{bmatrix} 3 \\ 5 \end{bmatrix}$.

a. 求 A^{-1} 且用它解下列四个方程组

$$A\boldsymbol{x}=\boldsymbol{b}_1,\ A\boldsymbol{x}=\boldsymbol{b}_2,\ A\boldsymbol{x}=\boldsymbol{b}_3,\ A\boldsymbol{x}=\boldsymbol{b}_4$$

b. （a）中的四个方程组可利用同样的行变换求解，因系数矩阵是相同的. 利用对增广矩阵 $[A\ \ \boldsymbol{b}_1\ \ \boldsymbol{b}_2\ \ \boldsymbol{b}_3\ \ \boldsymbol{b}_4]$ 做行变换的方法，解（a）中的四个方程组.

8. 设 P 可逆，$A=PBP^{-1}$，用 A 表示 B.

习题 9 和习题 10 中，标出命题的真假，给出理由.

9. a. 为了使矩阵 B 为 A 的逆，$AB=I$ 及 $BA=I$ 都必须为真.

b. 若 A,B 是可逆 $n\times n$ 矩阵，则 $A^{-1}B^{-1}$ 是 AB 的逆.

c. 若 $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ 且 $ab-cd\neq 0$ 则 A 可逆.

d. 若 A 是可逆 $n\times n$ 矩阵，则方程 $A\boldsymbol{x}=\boldsymbol{b}$ 对 $\mathbb{R}^n$ 中任意 $\boldsymbol{b}$ 相容.

e. 每个初等矩阵都可逆.

10. a. 若 A 可逆，则把 A 化为单位矩阵的行变换将 A^{-1} 化为 I_n.

b. 若 A 可逆，则 A^{-1} 的逆就是 A 本身.

c. 若干个可逆 $n\times n$ 矩阵之积为可逆，且其逆为这些矩阵的逆按相同顺序的乘积.

d. 若 A 是 $n\times n$ 矩阵，且 $A\boldsymbol{x}=\boldsymbol{e}_j$ 对每个 $j\in\{1, 2, \cdots n,\}$ 都相容，则 A 是可逆的. 提示：$\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n$ 表示单位矩阵的列.

e. 若 A 可经行变换化为单位矩阵，则 A 可逆.

11. 设 A 为可逆 $n\times n$ 矩阵，B 为 $n\times p$ 矩阵，证明方程 $AX=B$ 有唯一解 $A^{-1}B$.

12. 利用矩阵代数证明：若 A 是可逆矩阵，矩阵 D 满足 $AD=I$，则 $D=A^{-1}$.

13. 设 $AB=AC$，B 与 C 为 $n\times p$ 矩阵，A 可逆，证明 $B=C$. 若 A 不可逆，是否仍有 $B=C$？

14. 设 $(B-C)D=0$, B 与 C 为 $m\times n$ 矩阵，D 可逆，证明 $B=C$.

15. 设 A 为可逆 $n\times n$ 矩阵，B 为 $n\times p$ 矩阵，解释为什么 $A^{-1}B$ 可由行化简求得.

若 $[A\ \ B]\sim\cdots\sim[I\ \ X]$ 则 $X=A^{-1}B$

若 A 是大于 2×2 的矩阵，则 $[A\ \ B]$ 的行化简比计算 A^{-1} 和 $A^{-1}B$ 要快得多.

16. 设 A,B 是 $n\times n$ 矩阵，B 可逆，AB 也可逆. 证明 A 可逆.（提示：令 $C=AB$，从此式求出 A.）
17. 设 A,B,C 为可逆 $n\times n$ 矩阵，找一个矩阵 D 满足 $(ABC)D=I$ 及 $D(ABC)=I$ 从而证明 ABC 也可逆.
18. 设 A,B,C 是方阵，B 可逆，解方程 $AB=BC$ 求 A.
19. 设 A,B,C 是 $n\times n$ 可逆矩阵，由方程 $C^{-1}(A+X)B^{-1}=I_n$ 求 X.
20. 设 A,B,X 是 $n\times n$ 矩阵，$A,X,A-AX$ 可逆，假设
$$(A-AX)^{-1}=X^{-1}B \tag{3}$$
    a. 说明为什么 B 是可逆的.
    b. 由（3）式求 X，如果需要对矩阵求逆，请说明为什么该矩阵是可逆的.
21. 说明若 $n\times n$ 矩阵 A 为可逆的，则它的各列为线性无关.
22. 说明若 $n\times n$ 矩阵 A 为可逆，则它的各列生成 $\mathbb{R}^n$.（提示：复习 1.4 节定理 4.）
23. 设 A 为 $n\times n$ 矩阵，方程 $A\mathbf{x}=\mathbf{0}$ 仅有平凡解，说明为什么 A 有 n 个主元列和行等价于 I_n. 由定理 7，这说明 A 必定可逆（本题与 24 题将在 2.3 节用到）.
24. 设对 $n\times n$ 矩阵 A，方程 $A\mathbf{x}=\mathbf{b}$ 对任意 $\mathbb{R}^n$ 中的 $\mathbf{b}$ 有解，说明 A 必为可逆.（提示：A 是否行等价于 I_n？）

习题 25 和习题 26 对 $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ 证明定理 4.

25. 若设 $ad-bc=0$，方程 $A\mathbf{x}=\mathbf{0}$ 有多于一个解，为什么这说明 A 不可逆？（提示：首先考虑 $a=b=0$. 其次，若 a,b 不全为 0，考虑向量 $\mathbf{x}=\begin{bmatrix} -b \\ a \end{bmatrix}$.）
26. 证明若 $ad-bc\neq 0$，则计算 A^{-1} 的公式成立.

习题 27 和习题 28 证明例 5 下面的框中有关初等矩阵的命题的特殊情况. 此处 A 是 3×3 矩阵，$I=I_3$.（一般的证明将需要更多的符号.）

27. a. 利用 2.1 节方程（1）证明对 $i=1,2,3$，有 $\text{row}_i(A)=\text{row}_i(I)\cdot A$.
    b. 证明：若交换 A 的 1 和 2 行，则结果可以写成 EA，其中 E 是由 I 交换 1 和 2 行所得的初等矩阵.
    c. 证明：若 A 的行 3 乘以 5，则结果可以写成 EA，其中 E 是由 I 的行 3 乘以 5 所得的初等矩阵.
28. 证明：若 A 的行 2 被换成 $\text{row}_2(A)-3\cdot\text{row}_1(A)$，则结果可以写成 EA，其中 E 是由 I 的行 2 被换成 $\text{row}_2(I)-3\cdot\text{row}_1(A)$ 所得的初等矩阵.

求出习题 29~32 的矩阵的逆，若它们存在的话. 使用本节介绍的算法.

29. $\begin{bmatrix} 1 & -3 \\ 4 & -9 \end{bmatrix}$
30. $\begin{bmatrix} 3 & 6 \\ 4 & 7 \end{bmatrix}$
31. $\begin{bmatrix} 1 & 0 & -2 \\ -3 & 1 & 4 \\ 2 & -3 & 4 \end{bmatrix}$
32. $\begin{bmatrix} 1 & 2 & -1 \\ -4 & -7 & 3 \\ -2 & -6 & 4 \end{bmatrix}$
33. 利用本节的算法求矩阵 $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ 和 $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$ 的逆. 设 A 是相应的 $n\times n$ 矩阵，B 是它的逆，猜想 B 的形式，证明 $AB=I$.
34. 重复 33 题的方法猜想 $A=\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 2 & 2 & 0 & & 0 \\ 3 & 3 & 3 & & 0 \\ \vdots & & & \ddots & \vdots \\ n & n & n & \cdots & n \end{bmatrix}$ 的逆 B，证明 $AB=I$.

35. 设 $A=\begin{bmatrix}-1 & -7 & -3\\ 2 & 15 & 6\\ 1 & 3 & 2\end{bmatrix}$，求出 A^{-1} 的第 2 列和第 3 列而不计算第 1 列.

36. [M]设 $A=\begin{bmatrix}-25 & -9 & -27\\ 536 & 185 & 537\\ 154 & 52 & 143\end{bmatrix}$，求出 A^{-1} 的第 2 列和第 3 列而不计算第 1 列.

37. 设 $A=\begin{bmatrix}1 & 2\\ 1 & 3\\ 1 & 5\end{bmatrix}$，只使用 1，−1 和 0 作为元素（通过试错）构造一个 2×3 矩阵 C，使得 $CA=I_2$，计算 AC 并使 $AC\neq I_3$.

38. 设 $A=\begin{bmatrix}1 & -1 & 1 & 0\\ 0 & 1 & -1 & 1\end{bmatrix}$，只使用 1 和 0 作为元素构造一个 4×2 矩阵 D，使得 $AD=I_2$. 是否存在 4×2 矩阵 C，使得 $CA=I_4$？为什么？

39. [M]设 $D=\begin{bmatrix}0.011 & 0.003 & 0.001\\ 0.003 & 0.009 & 0.003\\ 0.001 & 0.003 & 0.011\end{bmatrix}$ 是一个弹性矩阵，弹性的单位是英寸/磅. 设在例 3 的图 2-5 中，在点 1，2，3 分别有 40，50 和 30 磅的力作用，求相应的形变.

40. [M]计算习题 39 中 D 的刚性矩阵 D^{-1}，求出在点 1 处引起 0.04 英寸形变所需要的作用力，设其他两点处的形变为 0.

41. [M]设 $D=\begin{bmatrix}0.0130 & 0.0050 & 0.0020 & 0.0010\\ 0.0050 & 0.0100 & 0.0040 & 0.0020\\ 0.0020 & 0.0040 & 0.0100 & 0.0050\\ 0.0010 & 0.0020 & 0.0050 & 0.0130\end{bmatrix}$ 为一条有四个受力点的弹性梁的弹性矩阵，单位为厘米/牛顿，在四个受力点处测得形变分别为 0.07、0.12、0.16 与 0.12 厘米，确定在这四个点上的作用力.

42. [M]对习题 41 中的 D，确定一组力，它在第 2 个点上引起形变 0.22 厘米，在其他三个点上引起的形变为 0. 这个答案与 D^{-1} 的元素有什么关系？（提示：先考虑在第 2 个点上引起 1 厘米形变的问题.）

练习题答案

1. a. $\det\begin{bmatrix}3 & -9\\ 2 & 6\end{bmatrix}=3\times6-(-9)\times2=18+18=36$，行列式不等于零，矩阵可逆.

   b. $\det\begin{bmatrix}4 & -9\\ 0 & 5\end{bmatrix}=4\times5-(-9)\times0=20\neq0$，矩阵可逆.

   c. $\det\begin{bmatrix}6 & -9\\ -4 & 6\end{bmatrix}=6\times6-(-9)\times(-4)=36-36=0$，矩阵不可逆.

2. $[A\ \ I]\sim\begin{bmatrix}1 & -2 & -1 & 1 & 0 & 0\\ -1 & 5 & 6 & 0 & 1 & 0\\ 5 & -4 & 5 & 0 & 0 & 1\end{bmatrix}\sim\begin{bmatrix}1 & -2 & -1 & 1 & 0 & 0\\ 0 & 3 & 5 & 1 & 1 & 0\\ 0 & 6 & 10 & -5 & 0 & 1\end{bmatrix}\sim\begin{bmatrix}1 & -2 & -1 & 1 & 0 & 0\\ 0 & 3 & 5 & 1 & 1 & 0\\ 0 & 0 & 0 & -7 & -2 & 1\end{bmatrix}$

   我们得到形如 $[B\ \ D]$ 的矩阵，其中 B 是方阵，有一个零行，进一步的行变换不可能将 B 变为 I，因此我们停止，A 没有逆.

2.3　可逆矩阵的特征

本节复习第 1 章引入的大部分重要概念，并且与 n 个未知量 n 个方程的方程组，以及方阵联系起来，主要结论是定理 8.

定理 8（可逆矩阵定理）

设 A 为 $n\times n$ 矩阵，则下列命题是等价的，即对某一特定的 A，它们同时为真或同时为假.

a. A 是可逆矩阵.

b. A 等价于 $n\times n$ 单位矩阵.

c. A 有 n 个主元位置.

d. 方程 $A\boldsymbol{x}=\boldsymbol{0}$ 仅有平凡解.

e. A 的各列线性无关.

f. 线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 是一对一的.

g. 对 $\mathbb{R}^n$ 中任意 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 至少有一个解.

h. A 的各列生成 $\mathbb{R}^n$.

i. 线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 把 $\mathbb{R}^n$ 映上到 $\mathbb{R}^n$ 上.

j. 存在 $n\times n$ 矩阵 C 使 $CA=I$.

k. 存在 $n\times n$ 矩阵 D 使 $AD=I$.

l. A^{T} 是可逆矩阵.

首先，我们需要某些记号，若当命题（a）为真则（j）也真，我们称（a）蕴涵（j），记为（a）⇒（j）. 我们将按图 2-6 中蕴涵的“循环”来证明这些命题的等价性，即这五个命题之一为真可推出其他命题也真，然后我们将把其他命题链接进这个循环.

图　2-6

证　若（a）为真，则 A^{-1} 可作为（j）中的 C，故（a）⇒（j），其次，由 2.1 节 23 题（请参阅该习题），(j）⇒（d），又由 2.2 节 23 题可知（d）⇒（c）. 若 A 是方阵且有 n 个主元位置，则主元必定在主对角线上，在这种情况下，A 的简化阶梯形是 I_n，因此（c）⇒（b）. 同时由 2.2 节定理 7 知（b）⇒（a）. 至此完成图 2-6 中的证明循环.

其次，由于 A^{-1} 可作为 D，(a）⇒（k）. 又由 2.1 节习题 26 知（k）⇒（g），而由 2.2 节习题 24 有（g）⇒（a），因此（g）和（k）被链接进这个循环. 再根据 1.4 节定理 4 和 1.9 节定理 12（a），得到对任一矩阵来说，(g）、(h）和（i）是等价的，因此，通过（g）使（h）和（i）被链接进这个循环.

因（d）、(e）、(f）对任一矩阵 A 是等价的（参见 1.7 节及 1.9 节定理 12（b）)，而（d）在这个循环之中，所以（e）和（f）也在这个循环中. 最后，由 2.2 节定理 6（c）有（a）⇒（l)，再根据同一个定理，将 A 和 A^{T} 互换后得到（l）⇒（a）. 见图 2-7. 这就完成了定理 8 的证明. ■

由 2.2 节定理 5，定理 8 中命题（g）也可写成“方程 $A\boldsymbol{x}=\boldsymbol{b}$ 对任意 $\mathbb{R}^n$ 中的 $\boldsymbol{b}$ 有唯一解”. 这

命题当然也蕴涵（b），因此也蕴涵 A 为可逆阵.

$$\begin{array}{ccc} & (k) & \\ \nearrow & & \searrow \\ (a) & \Leftarrow & (g) \end{array}$$

$$(g) \Leftrightarrow (h) \Leftrightarrow (i)$$

$$(d) \Leftrightarrow (e) \Leftrightarrow (f)$$

$$(a) \Leftrightarrow (l)$$

图 2-7

下列事实由定理 8 及 2.2 节习题 12 推出.

> 设 A 和 B 为方阵，若 $AB = I$，则 A 和 B 都是可逆的，且 $B = A^{-1}$，$A = B^{-1}$.

可逆矩阵定理将所有 $n \times n$ 矩阵分为两个不相交集合：可逆（非奇异）矩阵和不可逆（奇异）矩阵. 定理中每个命题给出了 $n \times n$ 可逆矩阵的一个性质. 定理中每个命题的否命题给出了 $n \times n$ 奇异矩阵的一个性质. 例如，每个 $n \times n$ 奇异矩阵不行等价于 I_n，没有 n 个主元位置，它的各列线性相关，其他的否命题在习题中考虑.

例 1 应用可逆矩阵定理来判断 A 是否可逆：

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & -2 \\ -5 & -1 & 9 \end{bmatrix}$$

解

$$A \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \\ 0 & -1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{bmatrix}$$

所以 A 有 3 个主元位置，根据可逆矩阵定理命题（c），A 是可逆的. ■

可逆矩阵定理的作用在于它给出了许多重要概念的联系，例如矩阵 A 的列的线性无关性与形如 $A\boldsymbol{x} = \boldsymbol{b}$ 的解的存在性关联起来. 但是必须强调，可逆矩阵定理仅能用于方阵. 例如，若一个 4×3 矩阵的列线性无关，我们不能用可逆矩阵定理断定形如 $A\boldsymbol{x} = \boldsymbol{b}$ 的方程的解的存在性或不存在性.

可逆线性变换

回忆 2.1 节矩阵乘法对应于线性变换的复合. 当矩阵 A 可逆时，方程 $A^{-1}A\boldsymbol{x} = \boldsymbol{x}$ 可看作关于线性变换的一个命题，见图 2-8.

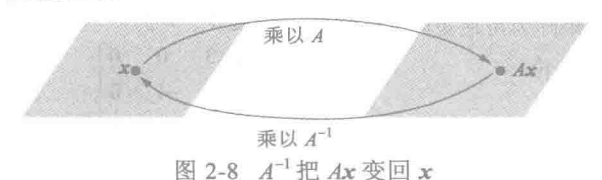


图 2-8 A^{-1} 把 $A\boldsymbol{x}$ 变回 $\boldsymbol{x}$

线性变换 $T:\mathbb{R}^n \to \mathbb{R}^n$ 称为**可逆的**，若存在函数 $S:\mathbb{R}^n \to \mathbb{R}^n$ 使得

$$\text{对所有 } \mathbb{R}^n \text{ 中的 } \boldsymbol{x}, \quad S(T(\boldsymbol{x})) = \boldsymbol{x} \tag{1}$$

$$\text{对所有 } \mathbb{R}^n \text{ 中的 } \boldsymbol{x}, \quad T(S(\boldsymbol{x})) = \boldsymbol{x} \tag{2}$$

下列定理说明若这样的 S 存在，它是唯一的而且必是线性变换. 我们称 S 是 T 的**逆**，把它写成 T^{-1}.

定理 9 设 $T:\mathbb{R}^n \to \mathbb{R}^n$ 为线性变换，A 为 T 的标准矩阵. 则 T 可逆当且仅当 A 是可逆矩阵. 这时由 $S(\boldsymbol{x}) = A^{-1}\boldsymbol{x}$ 定义的线性变换 S 是满足（1）和（2）的唯一函数.

证 设 T 是可逆的，则（2）说明 T 是从 $\mathbb{R}^n$ 映上到 $\mathbb{R}^n$ 的映射，因若 $\boldsymbol{b}$ 属于 $\mathbb{R}^n$，$\boldsymbol{x} = S(\boldsymbol{b})$，则 $T(\boldsymbol{x}) = T(S(\boldsymbol{b})) = \boldsymbol{b}$，所以每个 $\boldsymbol{b}$ 属于 T 的值域，于是由可逆矩阵定理命题（i），A 为可逆的.

反之，若 A 是可逆的，令 $S(\boldsymbol{x}) = A^{-1}\boldsymbol{x}$，则 S 是线性变换，且显然 S 满足（1）和（2），例如

$$S(T(\boldsymbol{x})) = S(A\boldsymbol{x}) = A^{-1}(A\boldsymbol{x}) = \boldsymbol{x}$$

于是 T 是可逆的. S 的唯一性的证明见习题 38. ■

例 2 设 $T:\mathbb{R}^n \to \mathbb{R}^n$ 是一对一线性变换，则 T 会如何？

解 T 的标准矩阵 A 的列是线性无关的（依 1.9 节定理 12），所以依可逆矩阵定理，A 是可逆的，而且 T 把 $\mathbb{R}^n$ 映上到 $\mathbb{R}^n$. 同时，依定理 9，T 为可逆. ■

数值计算的注解 实际工作中，你将会遇到"接近奇异的"或者**病态矩阵**——一个可逆矩阵，但当它的某些元素稍微改变就变成奇异矩阵. 在这种情况下，行变换可能由于舍入误差产生少于 n 个主元位置. 另外，有时舍入误差也可能使奇异矩阵变成是可逆的.

某些矩阵程序对一个方阵计算它的**条件数**，条件数越大，矩阵越接近奇异. 单位矩阵的条件数是 1，奇异矩阵的条件数为无穷大. 在极端情况下，矩阵程序可能无法区别奇异矩阵与病态矩阵.

习题 41~45 说明当条件数大时，矩阵计算可能产生明显的错误.

练习题

1. 确定 $A = \begin{bmatrix} 2 & 3 & 4 \\ 2 & 3 & 4 \\ 2 & 3 & 4 \end{bmatrix}$ 是否可逆.
2. 设对某个矩阵 A，可逆矩阵定理命题（g）不成立. 那么形如 $A\boldsymbol{x} = \boldsymbol{b}$ 的方程会如何？
3. 设 A, B 是 $n \times n$ 矩阵，方程 $AB\boldsymbol{x} = \boldsymbol{0}$ 有非平凡解，那么矩阵 AB 会如何？

习题 2.3

除非另有说明，本习题中的矩阵都是 $n \times n$ 矩阵，确定习题 1~10 中哪些矩阵为可逆矩阵. 使用尽可能少的计算，检验你的结论.

1. $\begin{bmatrix} 5 & 7 \\ -3 & -6 \end{bmatrix}$

2. $\begin{bmatrix} -4 & 2 \\ 6 & -3 \end{bmatrix}$

3. $\begin{bmatrix} 3 & 0 & 0 \\ -3 & -4 & 0 \\ 8 & 5 & -3 \end{bmatrix}$

4. $\begin{bmatrix} -5 & 1 & 4 \\ 0 & 0 & 0 \\ 1 & 4 & 9 \end{bmatrix}$

5. $\begin{bmatrix} 3 & 0 & -3 \\ 2 & 0 & 4 \\ -4 & 0 & 7 \end{bmatrix}$

6. $\begin{bmatrix} 1 & -3 & -6 \\ 0 & 4 & 3 \\ -3 & 6 & 0 \end{bmatrix}$

7. $\begin{bmatrix} -1 & -3 & 0 & 1 \\ 3 & 5 & 8 & -3 \\ -2 & -6 & 3 & 2 \\ 0 & -1 & 2 & 1 \end{bmatrix}$

8. $\begin{bmatrix} 3 & 4 & 7 & 4 \\ 0 & 1 & 4 & 6 \\ 0 & 0 & 2 & 8 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

9. [M] $\begin{bmatrix} 4 & 0 & -3 & -7 \\ -6 & 9 & 9 & 9 \\ 7 & -5 & 10 & 19 \\ -1 & 2 & 4 & -1 \end{bmatrix}$

10. [M] $\begin{bmatrix} 5 & 3 & 1 & 7 & 9 \\ 6 & 4 & 2 & 8 & -8 \\ 7 & 5 & 3 & 10 & 9 \\ 9 & 6 & 4 & -9 & -5 \\ 8 & 5 & 2 & 11 & 4 \end{bmatrix}$

习题 11~12 中，矩阵都是 $n\times n$，习题的每个部分都是形如"若<命题 1>则<命题 2>"的蕴涵式，如果当<命题 1>为真时<命题 2>总是为真，标记该蕴涵式为真．如果有一个例子给出<命题 2>为假但<命题 1>为真，则标记该蕴涵式为假．验证你的结论．

11. a. 若方程 $A\mathbf{x}=\mathbf{0}$ 仅有平凡解，则 A 行等价于 $n\times n$ 单位矩阵．
    b. 若 A 的各列生成 $\mathbb{R}^n$，则它的列线性无关．
    c. 若 A 是 $n\times n$ 矩阵，则对 $\mathbb{R}^n$ 中每个 $\mathbf{b}$，方程 $A\mathbf{x}=\mathbf{b}$ 至少有一个解．
    d. 若方程 $A\mathbf{x}=\mathbf{0}$ 有非平凡解，则 A 的主元位置少于 n 个．
    e. 若 A^{T} 不可逆，则 A 也不可逆．
12. a. 若存在 $n\times n$ 矩阵 D 使得 $AD=I$，则 $DA=I$．
    b. 若线性变换 $\mathbf{x}\mapsto A\mathbf{x}$ 把 $\mathbb{R}^n$ 映到 $\mathbb{R}^n$ 内，那么 A 行等价于 I．
    c. 若 A 的各列线性无关，则 A 的各列生成 $\mathbb{R}^n$．
    d. 若对每个 $\mathbb{R}^n$ 中的 $\mathbf{b}$，方程 $A\mathbf{x}=\mathbf{b}$ 至少有一个解，则变换不是一对一的．
    e. 若存在 $\mathbb{R}^n$ 中的 $\mathbf{b}$，方程 $A\mathbf{x}=\mathbf{b}$ 相容，则解是唯一的．
13. 若一个 $m\times n$ 矩阵的主对角线以下元素全为 0，则称之为**上三角矩阵**（如习题 8）．什么时候一个上三角矩阵是可逆的？验证你的答案．
14. 若一个 $m\times n$ 矩阵的主对角线以上元素全为 0，则称之为**下三角矩阵**（如习题 3）．什么时候一个下三角矩阵是可逆的？验证你的答案．
15. 一个 4×4 矩阵的各列不生成 $\mathbb{R}^4$，它是否可能可逆？为什么？
16. 若 $n\times n$ 矩阵 A 可逆，则 A^{T} 的各列线性无关，说明为什么？
17. 有两列相同的方阵是否可逆？为什么？
18. 有两行相同的方阵是否可逆？为什么？
19. 若 7×7 矩阵 D 的各列线性无关，关于方程 $D\mathbf{x}=\mathbf{b}$ 的解会如何？为什么？
20. 若 A 是 5×5 矩阵，对 $\mathbb{R}^5$ 中的每一 $\mathbf{b}$，方程 $A\mathbf{x}=\mathbf{b}$ 相容，是否可能对某个 $\mathbf{b}$，方程 $A\mathbf{x}=\mathbf{b}$ 有多个解？为什么？
21. 若方程 $C\mathbf{u}=\mathbf{v}$ 对 $\mathbb{R}^n$ 中某个 $\mathbf{v}$ 有多个解，$n\times n$ 矩阵 C 的各列能否生成 $\mathbb{R}^n$？为什么？
22. 若 $n\times n$ 矩阵 E和F 满足性质 $EF=I$，则 E和F 是可交换的．说明这是为什么．
23. 设 F 是 $n\times n$ 矩阵．若方程 $F\mathbf{x}=\mathbf{y}$ 对 $\mathbb{R}^n$ 中的某个 $\mathbf{y}$ 不相容，方程 $F\mathbf{x}=\mathbf{0}$ 会如何？为什么？
24. 若 $n\times n$ 矩阵 G 不能行等价于 I_n，G 的列会如何？为什么？
25. 验证例 1 前面框内的命题．
26. 说明为什么当 A 的各列线性无关时，A^2 的各

列可以生成 $\mathbb{R}^n$.

27. 设 A 和 B 是 $n\times n$ 矩阵. 证明：若 AB 可逆，则 A 也可逆. 你不能使用定理 6（b），因为你不能假定 A 和 B 是可逆的.（提示：存在一个矩阵 W 使得 $ABW=I$，为什么？）
28. 设 A 和 B 是 $n\times n$ 矩阵. 证明：若 AB 可逆，则 B 也可逆.
29. 若 A 是 $n\times n$ 矩阵，变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 是一对一的，其他关于这个变换的看法会如何？验证你的答案.
30. 若 A 是 $n\times n$ 矩阵，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 对 $\mathbb{R}^n$ 中的某些 $\boldsymbol{b}$ 有多个解，则变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 不是一对一的. 其他关于这个变换的看法会如何？验证你的答案.
31. 设 A 是 $n\times n$ 矩阵，对 $\mathbb{R}^n$ 中的每个 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 至少有一个解，不用定理 5 或 8，说明为什么每个方程 $A\boldsymbol{x}=\boldsymbol{b}$ 事实上恰好有一个解.
32. 设 A 是 $n\times n$ 矩阵，方程 $A\boldsymbol{x}=\boldsymbol{0}$ 仅有平凡解. 不用可逆矩阵定理，说明对 $\mathbb{R}^n$ 中的每个 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 必有一个解.

习题 33 和习题 34 中，T 是由 $\mathbb{R}^2$ 到 $\mathbb{R}^2$ 内的线性变换，说明 T 可逆并求出 T^{-1}.

33. $T(x_1,x_2)=(-5x_1+9x_2,4x_1-7x_2)$
34. $T(x_1,x_2)=(2x_1-8x_2,-2x_1+7x_2)$
35. 设 $T:\mathbb{R}^n\to\mathbb{R}^n$ 为可逆线性变换，说明为什么 T 既是一对一的又是映上到 $\mathbb{R}^n$ 的. 利用方程（1）和（2），运用一个或多个定理给出第二种解释.
36. 设 $T:\mathbb{R}^n\to\mathbb{R}^n$ 为线性变换，对 $\mathbb{R}^n$ 中一对不同的 $\boldsymbol{u}$ 和 $\boldsymbol{v}$，有 $T(\boldsymbol{u})=T(\boldsymbol{v})$，$T$ 能否将 $\mathbb{R}^n$ 映上到 $\mathbb{R}^n$？为什么？
37. 设 T和U 是 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 的线性变换，对 $\mathbb{R}^n$ 中的所有 $\boldsymbol{x}$，有 $T(U(\boldsymbol{x}))=\boldsymbol{x}$. 对 $\mathbb{R}^n$ 中的所有 $\boldsymbol{x}$，$U(T(\boldsymbol{x}))=\boldsymbol{x}$ 是否成立？为什么？
38. 设 $T:\mathbb{R}^n\to\mathbb{R}^n$ 为可逆线性变换，设 S 和 U 为 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 的函数，对一切 $\mathbb{R}^n$ 中的 $\boldsymbol{x}$，有 $S(T(\boldsymbol{x}))=\boldsymbol{x}$ 和 $U(T(\boldsymbol{x}))=\boldsymbol{x}$，证明对 $\mathbb{R}^n$ 中一切 $\boldsymbol{v}$，有 $U(\boldsymbol{v})=S(\boldsymbol{v})$. 这将证明 T 有唯一的逆，如定理 9 所说的那样.（提示：给定 $\mathbb{R}^n$ 中任意 $\boldsymbol{v}$，我们说对某个 $\boldsymbol{x}$ 有 $\boldsymbol{v}=T(\boldsymbol{x})$. 为什么？计算 $S(\boldsymbol{v})$和$U(\boldsymbol{v})$.）
39. 设 T 是将 $\mathbb{R}^n$ 映上到 $\mathbb{R}^n$ 的线性变换，证明 T^{-1} 存在且它将 $\mathbb{R}^n$ 映上到 $\mathbb{R}^n$，T^{-1} 是否是一对一的？
40. 设 T 和 S 满足可逆方程（1）和（2），其中 T 是线性变换. 直接证明 S 是线性变换.（提示：给定 $\mathbb{R}^n$ 中的 $\boldsymbol{u},\boldsymbol{v}$，设 $\boldsymbol{x}=S(\boldsymbol{u})$，$\boldsymbol{y}=S(\boldsymbol{v})$，则 $T(\boldsymbol{x})=\boldsymbol{u}$，$T(\boldsymbol{y})=\boldsymbol{v}$，为什么？把 S 作用于方程 $T(\boldsymbol{x})+T(\boldsymbol{y})=T(\boldsymbol{x}+\boldsymbol{y})$ 的两边. 同样地，证明 $T(c\boldsymbol{x})=cT(\boldsymbol{x})$.）
41. [M]设某一实验得出下列方程组：

$$\begin{aligned}4.5x_1+3.1x_2&=19.249\\1.6x_1+1.1x_2&=6.843\end{aligned}\qquad(3)$$

   a. 解方程组（3），同时解下面的方程组（4），它是由（3）的右边四舍五入到 2 位小数所得，在每种情形下，求出准确解.

$$\begin{aligned}4.5x_1+3.1x_2&=19.25\\1.6x_1+1.1x_2&=6.84\end{aligned}\qquad(4)$$

   b. （4）的各元素与（3）的对应元素的误差不超过 0.05%，求把（4）的解作为（3）的解的近似值时的相对误差.

   c. 用你的矩阵程序求出（3）中系数矩阵的条件数.

习题 42~44 说明，如何使用矩阵 A 的条件数来估计方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的计算解的精确度. 若 A 和 $\boldsymbol{b}$ 的元素大约精确到 r 位有效数字，而 A 的条件数约为 10^k（k 为正整数），则 $A\boldsymbol{x}=\boldsymbol{b}$ 的计算解大约精确到至少 $r-k$ 位有效数字.

42. [M]求出习题 9 中矩阵 A 的条件数，构造 $\mathbb{R}^4$ 中随机向量 $\boldsymbol{x}$，计算 $A\boldsymbol{x}=\boldsymbol{b}$，然后用你的矩阵程序计算方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的解 $\boldsymbol{x}_1$，$\boldsymbol{x}_1$ 和 $\boldsymbol{x}$ 有几位数字相同？找出你的矩阵程序准确存储的数字位数，用 $\boldsymbol{x}_1$ 代替准确解 $\boldsymbol{x}$ 时有多少位精确数字被

丢失？

43. [M]对习题 10 中的矩阵重复习题 42.

44. [M]对适当的 $\boldsymbol{b}$ 解 $A\boldsymbol{x}=\boldsymbol{b}$，以求得五阶希尔伯特（Hilbert）矩阵的逆的第 5 列.

$$A=\begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 & 1/5 \\ 1/2 & 1/3 & 1/4 & 1/5 & 1/6 \\ 1/3 & 1/4 & 1/5 & 1/6 & 1/7 \\ 1/4 & 1/5 & 1/6 & 1/7 & 1/8 \\ 1/5 & 1/6 & 1/7 & 1/8 & 1/9 \end{bmatrix}$$

你希望求出的解 $\boldsymbol{x}$ 的元素有多少位精确数字？请说明.（注：精确解为（630，−12 600，56 700，−88 200，44 100）.）

45. [M]某些矩阵程序，如 MATLAB，有命令可生成各阶希尔伯特矩阵. 若可能的话，用求逆命令求出 12 阶或更高阶的希尔伯特矩阵 A 的逆，计算 AA^{-1} 并报告你的结果.

练习题答案

1. A 的各列显然线性相关，因为第 2 列与第 3 列是第 1 列的倍数. 因此由可逆矩阵定理，A 不是可逆的.
2. 若（g）不成立，方程对 $\mathbb{R}^n$ 中至少一个 $\boldsymbol{b}$ 为不相容.
3. 应用可逆矩阵定理于矩阵 AB，假设 AB 可逆，则由命题（d）得到：$AB\boldsymbol{x}=\boldsymbol{0}$ 仅有平凡解，这与所给条件相矛盾，因此 AB 不是可逆的.

2.4 分块矩阵

我们既可以把矩阵看作一个数的矩形表，也可以把它看作一组列向量，后面这种看法起了很重要的作用，因而，我们想考虑 A 的其他**分块**，把它用水平线和竖直线分成几块，如下面例 1 所示. 分块矩阵也出现在线性代数的现代应用中，因为这些记号简化了许多讨论，并使矩阵计算中许多本质的结构显露出来. 本节也给出复习矩阵代数和可逆矩阵定理的机会.

例 1 矩阵

$$A=\left[\begin{array}{ccc|cc|c} 3 & 0 & -1 & 5 & 9 & -2 \\ -5 & 2 & 4 & 0 & -3 & 1 \\ \hline -8 & -6 & 3 & 1 & 7 & -4 \end{array}\right]$$

也可写成 2×3 **分块矩阵**

$$A=\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

的形状，它的元素是分块（或子矩阵）

$$A_{11}=\begin{bmatrix} 3 & 0 & -1 \\ -5 & 2 & 4 \end{bmatrix},\quad A_{12}=\begin{bmatrix} 5 & 9 \\ 0 & -3 \end{bmatrix},\quad A_{13}=\begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$$A_{21}=[-8\quad -6\quad 3],\quad A_{22}=[1\quad 7],\quad A_{23}=[-4]$$ ■

例 2 当某一矩阵 A 出现在物理问题的数学模型中时，例如，电子网络、传输系统或大公司等，会很自然地把 A 看作一个分块矩阵. 例如，若一个微型计算机电路板主要由 3 块超大规模的集成电路芯片组成，如图 2-9 所示，那么这电路板的矩阵可以写成一般形式

图 2-9

$$A=\left[\begin{array}{c|c|c}A_{11}&A_{12}&A_{13}\\\hline A_{21}&A_{22}&A_{23}\\\hline A_{31}&A_{32}&A_{33}\end{array}\right]$$

A 的“对角”线上的子矩阵，即 A_{11}, A_{22} 和 A_{33} 是有关 3 块超大规模集成电路本身的矩阵，而其他子矩阵则与这三块芯片之间的相互联系有关. ■

加法与标量乘法

若矩阵 A 与 B 有相同维数且被同样地分块，则自然矩阵的和 $A+B$ 也被同样地分块. 这时 $A+B$ 的每一块恰好是 A 和 B 对应分块的（矩阵）和. 分块矩阵乘以一个数也可以逐块计算.

分块矩阵的乘法

分块矩阵也可用通常的行列法则进行，就如每一块都是数一样，只要 A 的列的分法与 B 的行的分法一致.

例 3 设

$$A=\left[\begin{array}{ccc|cc}2&-3&1&0&-4\\1&5&-2&3&-1\\\hline 0&-4&-2&7&-1\end{array}\right]=\begin{bmatrix}A_{11}&A_{12}\\A_{21}&A_{22}\end{bmatrix},B=\left[\begin{array}{cc}6&4\\-2&1\\-3&7\\\hline -1&3\\5&2\end{array}\right]=\begin{bmatrix}B_1\\B_2\end{bmatrix}$$

A 的 5 列被分成 3 列一组和 2 列一组. B 的 5 行按同样方法分块——被分成 3 行一组和 2 行一组. 我们称 A 和 B 的分块是与分块乘法相一致的. AB 的乘积可以被写成

$$AB=\begin{bmatrix}A_{11}&A_{12}\\A_{21}&A_{22}\end{bmatrix}\begin{bmatrix}B_1\\B_2\end{bmatrix}=\begin{bmatrix}A_{11}B_1+A_{12}B_2\\A_{21}B_1+A_{22}B_2\end{bmatrix}=\left[\begin{array}{cc}-5&4\\-6&2\\\hline 2&1\end{array}\right]$$

重要的是，在 AB 的表达式中的小乘积，每一项应把来自 A 的子矩阵写在左边，因矩阵乘法是不可交换的. 例如

$$A_{11}B_1=\begin{bmatrix}2&-3&1\\1&5&-2\end{bmatrix}\begin{bmatrix}6&4\\-2&1\\-3&7\end{bmatrix}=\begin{bmatrix}15&12\\2&-5\end{bmatrix}$$

$$A_{12}B_2=\begin{bmatrix}0&-4\\3&-1\end{bmatrix}\begin{bmatrix}-1&3\\5&2\end{bmatrix}=\begin{bmatrix}-20&-8\\-8&7\end{bmatrix}$$

因此 AB 的上面一块是

$$A_{11}B_1+A_{12}B_2=\begin{bmatrix}15&12\\2&-5\end{bmatrix}+\begin{bmatrix}-20&-8\\-8&7\end{bmatrix}=\begin{bmatrix}-5&4\\-6&2\end{bmatrix}$$ ■

分块矩阵乘法的行列法则给出了两个矩阵乘积的最一般观点. 下面每一个有关矩阵乘积的观点已经使用简单的矩阵分块的思想讨论过：（1）使用 A 的列来给出 $A\boldsymbol{x}$ 的定义；（2）AB 的列的定义；（3）计算 AB 的行列法则；（4）A 的行与矩阵 B 的乘积作为 AB 的行. 在下面的定理

10 仍然应用分块的思想给出 AB 第 5 种观点.

下面的例子为定理 10 做准备. 符号 $\text{col}_k(A)$表示 A 的第 k 列，$\text{row}_k(B)$表示 B 的第 k 行.

例 4 设 $A=\begin{bmatrix}-3 & 1 & 2\\ 1 & -4 & 5\end{bmatrix}$和 $B=\begin{bmatrix}a & b\\ c & d\\ e & f\end{bmatrix}$. 证明

$$AB=\text{col}_1(A)\text{row}_1(B)+\text{col}_2(A)\text{row}_2(B)+\text{col}_3(A)\text{row}_3(B)$$

解 上面的每一项都是外积（见 2.1 节习题 27 和 28），由计算矩阵乘积的行列法则，有

$$\text{col}_1(A)\text{row}_1(B)=\begin{bmatrix}-3\\ 1\end{bmatrix}\begin{bmatrix}a & b\end{bmatrix}=\begin{bmatrix}-3a & -3b\\ a & b\end{bmatrix}$$

$$\text{col}_2(A)\text{row}_2(B)=\begin{bmatrix}1\\ -4\end{bmatrix}\begin{bmatrix}c & d\end{bmatrix}=\begin{bmatrix}c & d\\ -4c & -4d\end{bmatrix}$$

$$\text{col}_3(A)\text{row}_3(B)=\begin{bmatrix}2\\ 5\end{bmatrix}\begin{bmatrix}e & f\end{bmatrix}=\begin{bmatrix}2e & 2f\\ 5e & 5f\end{bmatrix}$$

于是

$$\sum_{k=1}^{3}\text{col}_k(A)\text{row}_k(B)=\begin{bmatrix}-3a+c+2e & -3b+d+2f\\ a-4c+5e & b-4d+5f\end{bmatrix}$$

这个矩阵恰好就是 AB. 注意 AB 的（1, 1）元素是 3 个外积的（1, 1）元素之和，AB 的（1, 2）元素是 3 个外积的（1, 2）元素之和，等等. ■

定理 10 （AB 的列行展开）

若 A 是 $m\times n$ 矩阵，B 是 $n\times p$ 矩阵，则

$$AB=\begin{bmatrix}\text{col}_1(A) & \text{col}_2(A) & \cdots & \text{col}_n(A)\end{bmatrix}\begin{bmatrix}\text{row}_1(B)\\ \text{row}_2(B)\\ \vdots\\ \text{row}_n(B)\end{bmatrix} \tag{1}$$

$$=\text{col}_1(A)\text{row}_1(B)+\cdots+\text{col}_n(A)\text{row}_n(B)$$

证 对每个行指标 i 和列指标 j，乘积 $\text{col}_k(A)\text{row}_k(B)$ 的 (i,j) 元素是 $\text{col}_k(A)$ 中元素 a_{ik} 与 $\text{row}_k(B)$ 中元素 b_{kj} 的积，因此在（1）的和中，(i,j) 元素为

$$\underset{(k=1)}{a_{i1}b_{1j}}+\underset{(k=2)}{a_{i2}b_{2j}}+\cdots+\underset{(k=n)}{a_{in}b_{nj}}$$

而根据行列法则，该和恰好是 AB 的 (i,j) 元素. ■

分块矩阵的逆

下例说明分块矩阵的逆的求法.

例 5 形如 $A=\begin{bmatrix}A_{11} & A_{22}\\ 0 & A_{22}\end{bmatrix}$的矩阵称为分块上三角矩阵，设 A_{11} 是 $p\times p$ 矩阵，A_{22} 是 $q\times q$ 矩阵，且 A 为可逆矩阵. 求 A^{-1} 的表达式.

解 用 B 表示 A^{-1} 且把它分块使

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}=\begin{bmatrix} I_p & 0 \\ 0 & I_q \end{bmatrix} \tag{2}$$

这个矩阵方程包含了4个有关未知子矩阵 $B_{11},\cdots,B_{22}$ 的方程，计算（2）式左边的乘积得

$$A_{11}B_{11}+A_{12}B_{21}=I_p \tag{3}$$

$$A_{11}B_{12}+A_{12}B_{22}=0 \tag{4}$$

$$A_{22}B_{21}=0 \tag{5}$$

$$A_{22}B_{22}=I_q \tag{6}$$

方程（6）本身并不能说明 A_{22} 可逆，因我们还不知道 $B_{22}A_{22}=I_q$，但应用可逆矩阵定理，及 A_{22} 是方阵的事实，可以断定 A_{22} 为可逆且 $B_{22}=A_{22}^{-1}$. 现在我们利用（5）式求得

$$B_{21}=A_{22}^{-1}0=0$$

因此（3）式简化为

$$A_{11}B_{11}+0=I_p$$

这说明 A_{11} 是可逆的，且 $B_{11}=A_{11}^{-1}$，最后由（4），

$$A_{11}B_{12}=-A_{12}B_{22}=-A_{12}A_{22}^{-1} \quad 和 \quad B_{12}=-A_{11}^{-1}A_{12}A_{22}^{-1}$$

于是

$$A^{-1}=\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}^{-1}=\begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix}$$ ■

分块对角矩阵是一个分块矩阵，除了主对角线上各分块外，其余全是零分块. 这样的一个矩阵是可逆的当且仅当主对角线上各分块都是可逆的. 见习题13和14.

数值计算的注解

1. 当矩阵太大时，不适于存储在高速计算机内存中，分块矩阵允许计算机一次处理两到三块子矩阵，例如，最近关于线性规划的工作中，一个研究团队把矩阵分为837行和51列以简化问题. 这个问题的解在Cray超计算机上大约需要4分钟[㊀].

2. 某些高速计算机，特别是具有向量传输技术的计算机，当把矩阵分块后再进行矩阵运算更有效[㊁].

3. 高性能数值计算的线性代数专业软件LAPACK，广泛使用分块矩阵进行计算.

下面的习题给出了运用矩阵代数的实践，表明了应用中的典型计算.

练习题

1. 证明 $\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}$ 可逆且求出它的逆.
2. 计算 $X^{\mathrm{T}}X$，其中 X 分块为 $[X_1 \quad X_2]$.

㊀ 若你不知道这51个分块的列每个包含大约250 000列，也许不觉得解题时间很短. 原来的问题有837个方程，包含12 750 000个变量. 矩阵中100亿个元素中大约有1亿个不等于0，参阅 Robert E. Bixby,et al. “Very Large-Scale Linear Programming: A Case Study in Combining Interior Point and Simplex Methods.” *Operations Research*, 40, no. 5（1992）: 885-897.

㊁ 分块矩阵算法对计算机的重要性见 *Matrix Computations*,3rd. ed., Gene H. Golub and Charles F. van Loan （Baltimore: Johns Hopkins University Press, 1996）.

习题 2.4

习题 1~9 中，假设这些矩阵的分块适于分块乘法，计算习题 1~4 的乘积.

1. $\begin{bmatrix} I & 0 \\ E & I \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}$

2. $\begin{bmatrix} E & 0 \\ 0 & F \end{bmatrix}\begin{bmatrix} P & Q \\ R & S \end{bmatrix}$

3. $\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}$

4. $\begin{bmatrix} I & 0 \\ -E & I \end{bmatrix}\begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$

习题 5~8 中，用 A,B,C 求出 X,Y,Z 的表达式，写出你的理由. 在这当中你需要对矩阵的维数做假设以得到一个公式.（提示：计算左边的乘积并使它等于右边.）

5. $\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}\begin{bmatrix} I & 0 \\ X & Y \end{bmatrix}=\begin{bmatrix} 0 & I \\ Z & 0 \end{bmatrix}$

6. $\begin{bmatrix} X & 0 \\ Y & Z \end{bmatrix}\begin{bmatrix} A & 0 \\ B & C \end{bmatrix}=\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

7. $\begin{bmatrix} X & 0 & 0 \\ Y & 0 & I \end{bmatrix}\begin{bmatrix} A & Z \\ 0 & 0 \\ B & I \end{bmatrix}=\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

8. $\begin{bmatrix} A & B \\ 0 & I \end{bmatrix}\begin{bmatrix} X & Y & Z \\ 0 & 0 & I \end{bmatrix}=\begin{bmatrix} I & 0 & 0 \\ 0 & 0 & I \end{bmatrix}$

9. 设 B_{11} 是可逆矩阵，求出矩阵 A_{21} 和 A_{31}（依据 B 的分块）使下列的积有所说的形式，并计算 C_{22}.（提示：计算左边的乘积使它等于右边.）

$$\begin{bmatrix} I & 0 & 0 \\ A_{21} & I & 0 \\ A_{31} & 0 & I \end{bmatrix}\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \\ B_{31} & B_{32} \end{bmatrix}=\begin{bmatrix} C_{11} & C_{12} \\ 0 & C_{22} \\ 0 & C_{32} \end{bmatrix}$$

10. 设 $\begin{bmatrix} I & 0 & 0 \\ A & I & 0 \\ B & D & I \end{bmatrix}$ 的逆为 $\begin{bmatrix} I & 0 & 0 \\ P & I & 0 \\ Q & R & I \end{bmatrix}$，求 P,Q,R.

习题 11~12 中，标出各个命题的真假. 验证你的结论.

11. a. 若 $A=[A_1 \quad A_2]$，$B=[B_1 \quad B_2]$，A_1 和 A_2 的维数分别与 B_1 和 B_2 的维数相同，则

$$A+B=[A_1+B_1 \quad A_2+B_2].$$

b. 若 $A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$，$B=\begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$，则 A 和 B 的分块适合于分块乘法.

12. a. 若 A_1,A_2,B_1 和 B_2 是 $n\times n$ 矩阵，$A=\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$，且 $B=[B_1 \quad B_2]$，则乘积 BA 有定义，但 AB 无定义.

b. 若 $A=\begin{bmatrix} P & Q \\ R & S \end{bmatrix}$，则 A 的转置为 $A^{\mathrm{T}}=\begin{bmatrix} P^{\mathrm{T}} & Q^{\mathrm{T}} \\ R^{\mathrm{T}} & S^{\mathrm{T}} \end{bmatrix}$.

13. 设 $A=\begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}$，其中 B 和 C 是方阵，证明 A 可逆当且仅当 B 和 C 都可逆.

14. 证明：例 5 中的分块上三角矩阵 A 可逆，当且仅当 A_{11} 和 A_{22} 都可逆.（提示：若 A_{11} 和 A_{22} 都可逆，例 5 中 A^{-1} 的表达式给出 A 的逆矩阵.）这个事实对许多估计矩阵特征值的计算机程序是很重要的，特征值在第 5 章讨论.

15. 当太空卫星发射之后，为使卫星在精确计算过的轨道上运行，需要校正它的位置. 如图 2-10 所示. 雷达屏幕给出一组向量 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k$，它们给出卫星在不同时间里的位置与计划轨道的比较. 设 X_k 表示矩阵 $[\boldsymbol{x}_1\cdots\boldsymbol{x}_k]$，矩阵 $G_k=X_kX_k^{\mathrm{T}}$ 需要在雷达分析数据时计算出来，当 $\boldsymbol{x}_{k+1}$ 到达时，新的 G_{k+1} 必须计算出来，因数据向量高速到达，所以计算负担很重. 分块矩阵的计算起很大作用. 计算 G_k 和 G_{k+1} 的列行展开，叙述从 G_k 如何计算 G_{k+1}.

16. 设 $A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$，若 A_{11} 可逆，则矩阵 $S=A_{22}-A_{21}A_{11}^{-1}A_{12}$ 称为 A_{11} 的舒尔补. 类似地，若 A_{22} 为可逆，则矩阵 $A_{11}-A_{12}A_{22}^{-1}A_{21}$ 称为 A_{22} 的舒尔补. 求出 X 与 Y 使

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}=\begin{bmatrix} I & 0 \\ X & I \end{bmatrix}\begin{bmatrix} A_{11} & 0 \\ 0 & S \end{bmatrix}\begin{bmatrix} I & Y \\ 0 & I \end{bmatrix} \tag{7}$$

17. 设（7）式左边的分块矩阵 A 为可逆，而 A_{11} 为可逆. 证明 A_{11} 的舒尔补 S 也是可逆的.（提示：（7）的右边的外面两个因子总是可逆的，证明这一点.）当 A 和 A_{11} 都可逆，（7）导出用 S^{-1}，A_{11}^{-1} 及 A 的其他元素计算 A^{-1} 的一个公式.

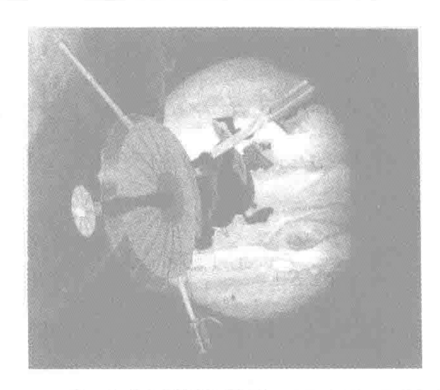


图 2-10 伽利略探测卫星于 1989 年 10 月 18 日发射，在 1995 年 11 月初到达接近木星的轨道

18. 设 X 是 $m\times n$ 矩阵，且 $X^{\mathrm{T}}X$ 为可逆，又设 $M=I_m-X(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}$. 加一列 $\boldsymbol{x}_0$ 于这组数据，构成矩阵 $W=[X\quad \boldsymbol{x}_0]$，计算 $W^{\mathrm{T}}W$. 它的（1,1）元素是 $X^{\mathrm{T}}X$，证明 $X^{\mathrm{T}}X$ 的舒尔补（16 题）是 $\boldsymbol{x}_0^{\mathrm{T}}M\boldsymbol{x}_0$，可以证明数 $(\boldsymbol{x}_0^{\mathrm{T}}M\boldsymbol{x}_0)^{-1}$ 是 $(W^{\mathrm{T}}W)^{-1}$ 的（2,2）元素，在适当假设下，这个数有一个有用的统计解释.

在研究物理系统的工程控制中，一组标准的微分方程用拉普拉斯变换转成下列线性方程组：

$$\begin{bmatrix} A-sI_n & B \\ C & I_m \end{bmatrix}\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{u} \end{bmatrix}=\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{y} \end{bmatrix} \qquad (8)$$

这里 A,B,C 分别为 $n\times n, n\times m, m\times n$ 矩阵. s 是一个变量. $\mathbb{R}^m$ 中向量 $\boldsymbol{u}$ 是系统的“输入”，$\mathbb{R}^m$ 中的向量 $\boldsymbol{y}$ 是“输出”，$\mathbb{R}^n$ 中向量 $\boldsymbol{x}$ 是“状态”向量.（实际上向量 $\boldsymbol{x},\boldsymbol{u},\boldsymbol{y}$ 都是 s 的函数，但忽略这一事实并不影响 19 题和 20 题的代数计算.）

19. 假设 $A-sI_n$ 是可逆的，把（8）看作两个矩阵方程的方程组，把第一个方程解的 $\boldsymbol{x}$ 代入第二个方程，结果是形如 $W(s)\boldsymbol{u}=\boldsymbol{y}$ 的方程，$W(s)$ 是依赖于 s 的矩阵，称为系统的传递函数，因它把输入 $\boldsymbol{u}$ 变换成输出 $\boldsymbol{y}$. 求出 $W(s)$，叙述它如何与（8）左边的分块的系统矩阵相关. 见 16 题.

20. 设 19 题中的传递函数 $W(s)$ 对某个 s 为可逆，可以证明，逆传递函数 $W(s)^{-1}$ 把输出转换为输入，是 $A-BC-sI_n$ 对于下列矩阵的舒尔补. 求出此舒尔补，见 16 题.

$$\begin{bmatrix} A-BC-sI_n & B \\ -C & I_m \end{bmatrix}$$

21. a. 证明 $A^2=I$，其中 $A=\begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}$

b. 利用分块矩阵证明 $M^2=I$，其中

$$M=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & -2 & 1 \end{bmatrix}$$

22. 推广习题 21 的思想，构造一个 6×6 矩阵 $M=\begin{bmatrix} A & 0 & 0 \\ 0 & B & 0 \\ C & 0 & D \end{bmatrix}$，使得 $M^2=I$，C 是 2×2 非零矩阵，验证你构造的矩阵满足要求.

23. 应用分块矩阵和归纳法证明两个下三角矩阵的乘积仍是下三角矩阵.（提示：一个 $(k+1)\times(k+1)$ 下三角矩阵 $A1$ 可以写成以下形式，a 是标量，$\boldsymbol{v}$ 是 $\mathbb{R}^k$ 中的向量，A 是 $k\times k$ 下三角矩阵，参见“学习指导”.）

$$A_1=\begin{bmatrix} a & \boldsymbol{0}^{\mathrm{T}} \\ \boldsymbol{v} & A \end{bmatrix}$$

24. 用分块矩阵归纳证明对 $n=2,3,\cdots$，下面的 $n\times n$ 矩阵 A 是可逆的，其逆为 B.

$$A=\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & & 0 \\ 1 & 1 & 1 & & 0 \\ \vdots & & & \ddots & \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix},\quad B=\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & & 0 \\ 0 & -1 & 1 & & 0 \\ \vdots & & \ddots & \ddots & \\ 0 & & \cdots & -1 & 1 \end{bmatrix}$$

归纳证明时，首先假设 A 和 B 是 $(k+1)\times(k+1)$ 矩阵，将 A 和 B 按类似习题 23 的方式进行分块.

25. 不使用行化简，求矩阵 $A=\begin{bmatrix} 1 & 2 & 0 & 0 & 0 \\ 3 & 5 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 7 & 8 \\ 0 & 0 & 0 & 5 & 6 \end{bmatrix}$ 的逆.

26. [M]对分块运算，必须能够访问或录入大矩阵的子矩阵. 叙述你的矩阵程序中实现下列功能

的命令．设 A 是 20×30 矩阵.

a. 显示 A 的从 5 行到 10 行，15 列到 20 列的子矩阵.

b. 把 5×10 矩阵 B 从第 5 行第 10 列开始插入 A.

c. 建立形如 $C=\begin{bmatrix} A & 0 \\ 0 & A^{\mathrm{T}} \end{bmatrix}$ 的 50×50 矩阵．（注：有可能不需要用命令说明 C 中的零分块矩阵.）

27. [M]设由于内存或维数限制，矩阵程序不可能存储多于 32 行 32 列的矩阵，假设某个项目涉及 50×50 矩阵 A 和 B，叙述你的矩阵程序中需要完成下列功能的命令.

a. 计算 $A+B$.

b. 计算 AB.

c. 对 $\mathbb{R}^{50}$ 中某个向量 $\boldsymbol{b}$，解方程 $A\boldsymbol{x}=\boldsymbol{b}$，假设 A 可分块为 2×2 分块矩阵 $[A_{ij}]$，其中 A_{11} 是可逆 20×20 矩阵，A_{22} 是可逆 30×30 矩阵，A_{12} 是零矩阵．（提示：用适当的小方程组来解，不使用矩阵的逆.）

练习题答案

1. 若 $\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}$ 可逆，它的逆有 $\begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$ 的形式．计算

$$\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}\begin{bmatrix} W & X \\ Y & Z \end{bmatrix}=\begin{bmatrix} W & X \\ AW+Y & AX+Z \end{bmatrix}$$

所以 W,X,Y,Z 满足 $W=I, X=0, AW+Y=0, AX+Z=I$．由此 $Y=-A$ 及 $Z=I$，因此

$$\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}\begin{bmatrix} I & 0 \\ -A & I \end{bmatrix}=\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

将左边两个矩阵对调，等式仍成立，所以所给分块矩阵为可逆阵，它的逆是 $\begin{bmatrix} I & 0 \\ -A & I \end{bmatrix}$（也可以用可逆矩阵定理）.

2. $X^{\mathrm{T}}X=\begin{bmatrix} X_1^{\mathrm{T}} \\ X_2^{\mathrm{T}} \end{bmatrix}[X_1, X_2]=\begin{bmatrix} X_1^{\mathrm{T}}X_1 & X_1^{\mathrm{T}}X_2 \\ X_2^{\mathrm{T}}X_1 & X_2^{\mathrm{T}}X_2 \end{bmatrix}$．$X^{\mathrm{T}}$ 的分块和 X 的分块是适于乘法的，因 X^{T} 的列数等于 X 的行数．$X^{\mathrm{T}}X$ 的分块常用于矩阵计算的计算机算法.

2.5 矩阵因式分解

矩阵 A 的因式分解是把 A 表示为两个或更多个矩阵的乘积，矩阵乘法是数据的综合（把两个或更多个线性变换的作用结合成一个矩阵），矩阵因式分解是数据的分解，在计算机科学的语言中，将 A 表示为矩阵的乘积是对 A 中数据的预处理，把这些数据组成两个或更多部分，这种结构可能更有用，或者更便于计算.

矩阵的因式分解，以及以后的线性变换的分解将在许多关键地方出现．本节主要讨论一种在许多重要的计算机程序核心部分出现的因式分解，某些其他的分解在习题中介绍.

LU 分解

下面所述的 LU 分解，是一些在工业与商业问题中常见的，解一系列具有相同系数矩阵的

线性方程（见 32 题）：

$$Ax = b_1, \qquad Ax = b_2, \cdots, Ax = b_p \tag{1}$$

在 5.8 节中，逆幂法通过逐个求解一系列形如（1）中的方程来估计矩阵的特征值.

当 A 为可逆，可计算 A^{-1}，然后计算 $A^{-1}b_1$，$A^{-1}b_2$，等等. 然而，在实践中，（1）中第一个方程是由行变换解出的，并同时得出 A 的 LU 分解，因而剩下的方程由 LU 分解求解.

首先，设 A 是 $m\times n$ 矩阵可以行化简为阶梯形而不必行对换（此后，我们将处理一般情形），则 A 可写成形式 $A=LU$，L 是 $m\times m$ 下三角矩阵，主对角线元素全是 1，U 是 A 的一个等价的 $m\times n$ 阶梯形矩阵. 例如，见图 2-11. 这样一个分解称为 LU **分解**，矩阵 L 是可逆的，称为单位下三角矩阵.

$$A=\underset{L}{\begin{bmatrix}1&0&0&0\\ *&1&0&0\\ *&*&1&0\\ *&*&*&1\end{bmatrix}}\underset{U}{\begin{bmatrix}\blacksquare&*&*&*&*\\ 0&\blacksquare&*&*&*\\ 0&0&0&\blacksquare&*\\ 0&0&0&0&0\end{bmatrix}}$$

图 2-11　LU 分解

在研究如何构造 L 和 U 之前，我们将看看它们为什么这么有用. 当 $A=LU$ 时，方程 $Ax=b$ 可写成 $L(Ux)=b$，把 Ux 写成 y，可以由解下面一对方程来求解 x：

$$\boxed{\begin{aligned}Ly&=b\\ Ux&=y\end{aligned}} \tag{2}$$

首先解 $Ly=b$ 然后解 $Ux=y$ 求得 x，见图 2-12. 每个方程都比较容易解，因 L 和 U 都是三角矩阵.

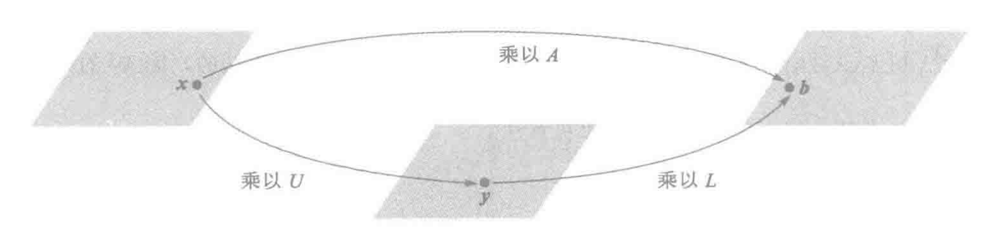


图 2-12　映射 $x\mapsto Ax$ 的分解

例 1　可以证明

$$A=\begin{bmatrix}3&-7&-2&2\\ -3&5&1&0\\ 6&-4&0&-5\\ -9&5&-5&12\end{bmatrix}=\begin{bmatrix}1&0&0&0\\ -1&1&0&0\\ 2&-5&1&0\\ -3&8&3&1\end{bmatrix}\begin{bmatrix}3&-7&-2&2\\ 0&-2&-1&2\\ 0&0&-1&1\\ 0&0&0&-1\end{bmatrix}=LU$$

应用 A 的 LU 分解来解 $Ax=b$，其中 $b=\begin{bmatrix}-9\\ 5\\ 7\\ 11\end{bmatrix}$.

解 解 $L\boldsymbol{y}=\boldsymbol{b}$ 仅需 6 次乘法和 6 次加法，因这些运算仅需对第 5 列进行.（在 L 的每个主元下的零会在行变换的选取中自动产生.）

$$[L\quad \boldsymbol{b}]=\begin{bmatrix}1&0&0&0&-9\\-1&1&0&0&5\\2&-5&1&0&7\\-3&8&3&1&11\end{bmatrix}\sim\begin{bmatrix}1&0&0&0&-9\\0&1&0&0&-4\\0&0&1&0&5\\0&0&0&1&1\end{bmatrix}=[I\quad \boldsymbol{y}]$$

对 $U\boldsymbol{x}=\boldsymbol{y}$，行化简的向后步骤需要 4 次除法，6 次乘法和 6 次加法.（例如，把 $[U\quad \boldsymbol{y}]$ 的第 4 列变成零需要 1 次除法和 3 次乘法——把第 4 行的倍数加到上面各行.）

$$[U\quad \boldsymbol{y}]=\begin{bmatrix}3&-7&-2&2&-9\\0&-2&-1&2&-4\\0&0&-1&1&5\\0&0&0&-1&1\end{bmatrix}\sim\begin{bmatrix}1&0&0&0&3\\0&1&0&0&4\\0&0&1&0&-6\\0&0&0&1&-1\end{bmatrix},\quad \boldsymbol{x}=\begin{bmatrix}3\\4\\-6\\-1\end{bmatrix}$$

为求 $\boldsymbol{x}$ 需 28 次算术运算，或“浮算”（浮点运算），不包括求 L 和 U 的运算在内. 相反，$[A\quad \boldsymbol{b}]$ 行化简成为 $[I\quad \boldsymbol{x}]$ 需 62 次运算. ■

LU 分解的计算效率依赖于如何求 L 和 U，下面的算法证明，把 A 行化简为阶梯形 U 的运算同时也求出 L 而基本上不需要额外的运算，因而也求出 LU 分解. 在第一次行化简后，L 和 U 就可用来解系数矩阵为 A 的额外的方程.

LU 分解算法

设 A 可以化为阶梯形 U. 化简过程中仅用行倍加变换，即把一行的倍数加于它下面的另一行. 这样，存在单位下三角初等矩阵 $E_1,\cdots,E_p$ 使

$$E_p\cdots E_1A=U \tag{3}$$

于是

$$A=(E_p\cdots E_1)^{-1}U=LU$$

其中

$$L=(E_p\cdots E_1)^{-1} \tag{4}$$

可以证明，单位下三角矩阵的逆也是单位下三角矩阵.（例如见 19 题.）于是 L 是单位下三角矩阵.

注意（3）中的行变换，它把 A 化为 U，所以也把（4）中的 L 化为 I，因

$$E_p\cdots E_1L=(E_p\cdots E_1)(E_p\cdots E_1)^{-1}=I$$

这一点是构造 L 的关键.

> **LU 分解的算法**
>
> 1. 如果可能的话，用一系列的倍加变换把 A 化为阶梯形.
> 2. 填充 L 的元素使相同的行变换把 L 变为 I.

第 1 步不是永远可能的，但当它可能时，上述讨论指出 LU 分解存在. 例 2 将说明如何实现第 2 步. 根据上面的说明，L 要满足使用与（3）中相同的 $E_1,\cdots,E_p$，有

$$(E_p\cdots E_1)L=I$$

于是，依可逆矩阵定理，L 是可逆的，$(E_p \cdots E_1) = L^{-1}$. 由（3），$L^{-1}A = U$，所以步骤 2 可求出所要的 L.

例 2 求下列矩阵的 LU 分解：

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix}$$

解 因 A 有 4 行，L 应为 4×4 矩阵，L 的第一列应该是 A 的第一列除以它的第一行主元元素：

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & & 1 & 0 \\ -3 & & & 1 \end{bmatrix}$$

比较 A 和 L 的第一列. 把 A 的第一列的后 3 个元素变成零的行变换同时也将 L 的后 3 个元素变成 0，同样的道理对 L 的其他各列也是成立的，让我们看一下 A 变成阶梯形 U 的过程. 即每个矩阵中标出的元素，这些元素用于确定把 A 变成 U 的行变换顺序.（见（5）中标出的元素.）

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix} = A_1 \tag{5}$$

$$\sim A_2 = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 & 7 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} = U$$

上式中标出的元素确定了将 A 化为 U 的行变换，在每个主元列，把标出的元素除以主元后将结果放入 L：

$$\begin{bmatrix} 2 \\ -4 \\ 2 \\ -6 \end{bmatrix} \quad \begin{bmatrix} 3 \\ -9 \\ 12 \end{bmatrix} \quad \begin{bmatrix} 2 \\ 4 \end{bmatrix} \quad [5]$$

$$\div 2 \qquad \div 3 \qquad \div 2 \qquad \div 5$$

$$\downarrow \qquad \downarrow \qquad \downarrow \qquad \downarrow$$

$$\begin{bmatrix} 1 & & & \\ -2 & 1 & & \\ 1 & -3 & 1 & \\ -3 & 4 & 2 & 1 \end{bmatrix}, \quad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & 2 & 1 \end{bmatrix}$$

容易证明，所求出的 L 和 U 满足 $LU = A$. ■

在实际工作中，行对换几乎总是必要的，因为部分主元法可以用来提高精确度.（回忆这种

算法总是选择一列中可以做为主元的元素中绝对值最大的一个作主元.）为了处理行对换，上述的LU分解算法可以稍做改变，以产生一个置换下三角矩阵L，就是说经过行的置换后它成为（单位）下三角矩阵. 所得的置换LU分解可与前面一样的途径解方程$A\mathbf{x}=\mathbf{b}$，只要在把$[L \quad \mathbf{b}]$化简为$[I \quad \mathbf{y}]$时按照L中主元的顺序从左到右进行，并从第一列主元开始. 介绍LU分解的参考书通常包含L为置换下三角矩阵的可能性. 详见学习指导书.

数值计算的注解 下列运算次数的计算适用于$n\times n$稠密矩阵A（大部分元素非零），n相当大，例如$n\geqslant 30$.㊀

1. 计算A的LU分解大约需要$2n^3/3$浮算（大约与把$[A \quad \mathbf{b}]$行化简的次数相同），而求A^{-1}大约需要$2n^3$浮算.

2. 解$L\mathbf{y}=\mathbf{b}$和$U\mathbf{x}=\mathbf{y}$大约需要$2n^2$浮算，因任意$n\times n$三角方程组可以用大约n^2浮算解出.

3. 把$\mathbf{b}$乘以A^{-1}也需要$2n^2$浮算，但结果可能不如由L和U得出的精确（由于计算A^{-1}及$A^{-1}\mathbf{b}$的舍入误差）.

4. 若A是稀疏矩阵（大部分元素为0），则L和U可能也是稀疏的，然而A^{-1}很可能是稠密的. 这时，用LU分解来解方程$A\mathbf{x}=\mathbf{b}$很可能比用A^{-1}快很多，见习题31.

电子工程中的矩阵因式分解

矩阵因式分解在构造具有某些性质的电子网络中碰到. 下列讨论仅是矩阵分解与电路设计联系的一个大概.

设图2-13中的方框表示某种电路，具有输入与输出. 用$\begin{bmatrix} v_1 \\ i_1 \end{bmatrix}$表示输入电压与电流（电压以伏特为单位，电流用安培为单位），输出电压与电流为$\begin{bmatrix} v_2 \\ i_2 \end{bmatrix}$，通常，变换$\begin{bmatrix} v_1 \\ i_1 \end{bmatrix}\mapsto\begin{bmatrix} v_2 \\ i_2 \end{bmatrix}$为线性的，也就是说，存在矩阵$A$，称为传递矩阵，使

$$\begin{bmatrix} v_2 \\ i_2 \end{bmatrix}=A\begin{bmatrix} v_1 \\ i_1 \end{bmatrix}$$

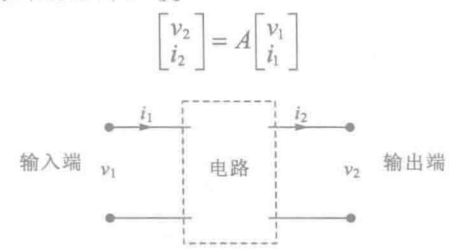


图2-13 具有输入端与输出端的电路

图2-14表示梯级网络，它有两个回路（也可以更多），它们串联起来，所以第一个电路的输出是第二个电路的输入. 图2-14左边的电路称为串联电路，具有电阻R_1（单位为欧姆），右端电路称为并联电路，有电阻R_2，应用欧姆定律，可以证明串联电路与并联电路的传递矩阵为

㊀ 见*Applied Linear Algebra*, 3rd ed. 3. 8节, Ben Noble and James W. Daniel（Englewood Cliffs, NJ: Prentice-Hall, 1988）. 记住，我们所说的浮算是指+，−，×，÷.

$$\begin{bmatrix} 1 & -R_1 \\ 0 & 1 \end{bmatrix} \qquad\qquad \begin{bmatrix} 1 & 0 \\ -1/R_2 & 1 \end{bmatrix}$$

串联电路的传递矩阵　　　　　　并联电路的传递矩阵

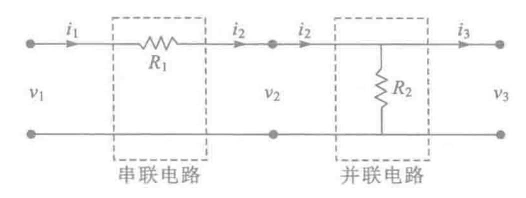


图 2-14　梯级电路

例 3

a. 计算图 2-14 中梯级网络的传递矩阵.

b. 设计一个传递矩阵为 $\begin{bmatrix} 1 & -8 \\ -0.5 & 5 \end{bmatrix}$ 的梯级网络.

解

a. 设 A_1 和 A_2 分别为串联电路与并联电路的传递矩阵，则输入向量 $\boldsymbol{x}$ 首先变换为 $A_1\boldsymbol{x}$，然后变为 $A_2(A_1\boldsymbol{x})$，两个电路的串联对应于变换的复合，所以梯级网络的传递矩阵为（注意顺序）

$$A_2A_1 = \begin{bmatrix} 1 & 0 \\ -1/R_2 & 1 \end{bmatrix}\begin{bmatrix} 1 & -R_1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1+R_1/R_2 \end{bmatrix} \tag{6}$$

b. 我们希望把矩阵 $\begin{bmatrix} 1 & -8 \\ -0.5 & 5 \end{bmatrix}$ 分解成两个传递矩阵的乘积，如（6）式，故我们要找出图 2-14 中 R_1 和 R_2 满足

$$\begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1+R_1/R_2 \end{bmatrix} = \begin{bmatrix} 1 & -8 \\ -0.5 & 5 \end{bmatrix}$$

由（1,2）元素，$R_1 = 8$ 欧姆，由（2,1）元素 $1/R_2 = 0.5$，而 $R_2 = 2$ 欧姆，对电阻的这些值，图 2-14 中网络有所需传递矩阵. ■

网络的传递函数总结了网络的输入一输出行为（设计规范），而不必去了解电路内部，为了物理上建立网络，使它有特定性质，工程师首先确定这样的网络是否可实现. 然后尝试把传递矩阵分解为对应于较小回路的传递矩阵的积，这些较小回路已经制造出来. 在交流电的情况下，传递矩阵的元素通常是有理复值函数（见 2.4 节习题 19 和 20，3.3 节例 2），标准问题是要找一个使用最少电子元件的最小实现.

练习题

求矩阵 $A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix}$ 的 LU 分解.

（注：A 仅有 3 个主元列，故例 2 的方法将仅产生 L 的前三列．L 的余下两列由 I_5 得到.）

习题 2.5

习题 1~6 中，用所给的 A 的 LU 分解来解方程 $A\boldsymbol{x}=\boldsymbol{b}$，在习题 1~2 中，同时用通常的行变换解方程 $A\boldsymbol{x}=\boldsymbol{b}$.

1. $A=\begin{bmatrix}3&-7&-2\\-3&5&1\\6&-4&0\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}-7\\5\\2\end{bmatrix}$

$A=\begin{bmatrix}1&0&0\\-1&1&0\\2&-5&1\end{bmatrix}\begin{bmatrix}3&-7&-2\\0&-2&-1\\0&0&-1\end{bmatrix}$

2. $A=\begin{bmatrix}2&-6&4\\-4&8&0\\0&-4&6\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}2\\-4\\6\end{bmatrix}$

$A=\begin{bmatrix}1&0&0\\-2&1&0\\0&1&1\end{bmatrix}\begin{bmatrix}2&-6&4\\0&-4&8\\0&0&-2\end{bmatrix}$

3. $A=\begin{bmatrix}2&-4&2\\-4&5&2\\6&-9&1\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}6\\0\\6\end{bmatrix}$

$A=\begin{bmatrix}1&0&0\\-2&1&0\\3&-1&1\end{bmatrix}\begin{bmatrix}2&-4&2\\0&-3&6\\0&0&1\end{bmatrix}$

4. $A=\begin{bmatrix}1&-1&2\\1&-3&1\\3&7&5\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}0\\-5\\7\end{bmatrix}$

$A=\begin{bmatrix}1&0&0\\1&1&0\\3&-5&1\end{bmatrix}\begin{bmatrix}1&-1&2\\0&-2&-1\\0&0&-6\end{bmatrix}$

5. $A=\begin{bmatrix}1&-2&-2&-3\\3&-9&0&-9\\-1&2&4&7\\-3&-6&26&2\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}1\\6\\0\\3\end{bmatrix}$

$A=\begin{bmatrix}1&0&0&0\\3&1&0&0\\-1&0&1&0\\-3&4&-2&1\end{bmatrix}\begin{bmatrix}1&-2&-2&-3\\0&-3&6&0\\0&0&2&4\\0&0&0&1\end{bmatrix}$

6. $A=\begin{bmatrix}1&3&4&0\\-2&-3&-4&12\\3&0&4&-36\\-5&-3&-8&49\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}1\\-2\\-1\\2\end{bmatrix}$

$A=\begin{bmatrix}1&0&0&0\\-2&1&0&0\\3&-3&1&0\\-5&4&-1&1\end{bmatrix}\begin{bmatrix}1&3&2&0\\0&3&0&12\\0&0&-2&0\\0&0&0&1\end{bmatrix}$

求习题 7~16 中矩阵 A 的 LU 分解（L 为单位下三角矩阵）. 注意 MATLAB 通常给出置换 LU 分解，因为它用部分主元法以提高精确度.

7. $\begin{bmatrix}2&5\\-3&-4\end{bmatrix}$

8. $\begin{bmatrix}6&9\\12&5\end{bmatrix}$

9. $\begin{bmatrix}3&-1&2\\-9&0&-4\\9&9&14\end{bmatrix}$

10. $\begin{bmatrix}-5&0&4\\10&2&-5\\10&10&16\end{bmatrix}$

11. $\begin{bmatrix}3&7&2\\6&19&4\\-3&-2&3\end{bmatrix}$

12. $\begin{bmatrix}2&3&2\\4&13&9\\-6&5&4\end{bmatrix}$

13. $\begin{bmatrix}1&3&-5&-3\\-1&-5&8&4\\4&2&-5&-7\\-2&-4&7&5\end{bmatrix}$

14. $\begin{bmatrix}1&3&1&5\\5&20&6&31\\-2&-1&1&-4\\-1&7&1&7\end{bmatrix}$

15. $\begin{bmatrix}2&0&5&2\\-6&3&-13&-3\\4&9&16&17\end{bmatrix}$

16. $\begin{bmatrix}2&-3&4\\-4&8&-7\\6&-5&14\\-6&9&-12\\8&-6&19\end{bmatrix}$

17. 当 A 可逆时，MATLAB 利用分解 $A=LU$（L 可能是置换下三角矩阵）来求 A^{-1}，即 $A^{-1}=U^{-1}L^{-1}$. 用这种方法求题 2 中 A 的逆（对 L 和 U 应用 2.2 节求逆的算法）.
18. 如 17 题，求习题 3 中 A 的逆.
19. 设 A 为下三角 $n\times n$ 矩阵，对角线上元素非零，证明 A 可逆且 A^{-1} 是下三角矩阵. [提示：说明为什么 A 可仅用倍加变换和倍乘变换变为 I.

（主元在何处？）接着说明为什么化简 A 的行变换把 A 变为 I 的同时把 I 变为下三角矩阵.]

20. 设 A 的 LU 分解为 $A=LU$. 说明为什么 A 可仅用行倍加变换成 U.（这里给的条件是课文中所证明的结论，而这里的结论是那里给出的条件.）

21. 设 $A=BC$，B 为可逆. 证明：任意一系列把 B 变成 I 的行变换也将 A 变成 C. 其逆不真，因零矩阵可分解为 $0=B\cdot 0$.

习题 22~26 介绍某些广泛应用的矩阵分解法，有些在本书中后面的章节讨论.

22. （简化 LU 分解）设 A 如练习题，求出 5×3 矩阵 B 及 3×4 矩阵 C，使 $A=BC$，推广这一想法到 A 是 $m\times n$ 矩阵的情形，$A=LU,U$ 仅有 3 个非零行.

23. （秩分解）设 $m\times n$ 矩阵 A 有分解 $A=CD$，C 是 $m\times 4$ 矩阵，D 是 $4\times n$ 矩阵.

    a. 证明 A 是 4 个外积的和.（见 2.4 节.）

    b. 设 $m=400$ 和 $n=100$，说明为什么程序员倾向于存储两个矩阵 C 和 D 而不存储矩阵 A.

24. （QR 分解）设 $A=QR$，其中 Q 和 R 都是 $n\times n$ 矩阵，R 可逆上三角矩阵，Q 满足 $Q^{\mathrm{T}}Q=I$，证明对任意属于 $\mathbb{R}^n$ 的 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 有唯一解，叙述求解的算法.

25. （奇异值分解）设 $A=UDV^{\mathrm{T}}$，这里 U,V 是 $n\times n$ 矩阵，$U^{\mathrm{T}}U=I,V^{\mathrm{T}}V=I$，$D$ 是对角矩阵，主对角线上元素 $\sigma_1,\sigma_2,\cdots,\sigma_n$ 为正数. 证明 A 可逆，并给出 A^{-1} 的一个表达式.

26. （谱分解）设 3×3 矩阵 A 可分解为 $A=PDP^{-1}$，P 是可逆 3×3 矩阵，D 是对角矩阵

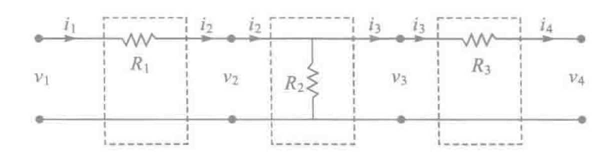
$$D=\begin{bmatrix}2&0&0\\0&3&0\\0&0&1\end{bmatrix}$$

证明在计算 A 的高次幂时，这种分解是有用的，使用 P 及 D 的元素求出 A^2,A^3 和 A^k（k 为正整数）的较为简单的公式.

27. 设计两个不同的梯级网络使其输入为 12 伏特与 6 安培而输出为 9 伏特与 4 安培.

28. 证明：若三个并联电路（电阻为 R_1,R_2,R_3）串联在一起，所得网络与一个单一的并联电路有相同的传递矩阵. 求该电路的电阻的表达式.

29. a. 计算下图的网络的传递矩阵.

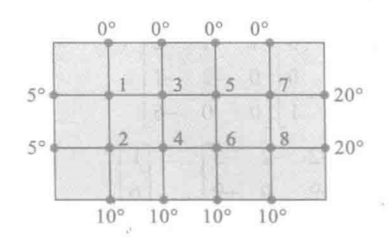


    b. 设 $A=\begin{bmatrix}3&-12\\-1/3&5/3\end{bmatrix}$，设计网络，使它的传递矩阵为 A，利用适当的矩阵分解.

30. 在习题 29 中，求出 A 的另一分解并用以设计传递矩阵为 A 的另一梯级网络.

31. [M]对下图中平板的稳态传热问题的解可近似地用方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的解来逼近，其中 $\boldsymbol{b}=(5,15,0,10,0,10,20,30)$，

$$A=\begin{bmatrix}4&-1&-1&&&&&\\-1&4&0&-1&&&&\\-1&0&4&-1&-1&&&\\&-1&-1&4&0&-1&&\\&&-1&0&4&-1&-1&\\&&&-1&-1&4&0&-1\\&&&&-1&0&4&-1\\&&&&&-1&-1&4\end{bmatrix}$$



（参阅 1.1 节习题 33.）A 中未写出的元素为 0. A 的非零元素都位于主对角线相邻的一条带中. 这样的带状矩阵在许多应用中出现，往往非常大（有数千的行和列但相对窄的带）.

a. 使用例 2 的方法求 A 的 LU 分解，注意两个因子

都是带状矩阵（有两条非零对角线在主对角线之上或之下）. 计算 $LU-A$ 来检验你的结果.

b. 使用 LU 分解解 $A\boldsymbol{x}=\boldsymbol{b}$.

c. 求出 A^{-1}，注意 A^{-1} 是稠密矩阵，无带状结构. 当 A 很大时，L 和 U 可以存储在比 A^{-1} 小得多的空间内. 这个事实是更多使用 A 的 LU 分解而不用 A^{-1} 本身的一个原因.

32. [M]下列带状矩阵可用来估计一根梁中的非稳态热传导，其中梁上的各点 $p_1,\cdots,p_4$ 的温度随时间变化. ㊀

Δx　Δx　Δx　Δx　Δx

p_1　p_2　p_3　p_4

矩阵中的常数 C 依赖于梁的物理性质，Δx 为各点之间距离，时间间隔 Δt 的长度是两次温度测量的间隔. 设对 $k=0,1,2,\cdots,\mathbb{R}^4$ 中向量 $\boldsymbol{t}_k$ 表示在 $k\Delta t$ 时刻各点的温度. 若梁的两端保持于 0°，则温度向量满足方程 $A\boldsymbol{t}_{k+1}=\boldsymbol{t}_k(k=0,1,\cdots)$，其中

$$A=\begin{bmatrix} (1+2C) & -C & & & \\ -C & (1+2C) & -C & & \\ & -C & (1+2C) & & \\ & & -C & -C & \\ & & & & (1+2C) \end{bmatrix}$$

a. 求出当 $C=1$ 时，A 的 LU 分解. 有三条非零对角线的矩阵称为三对角矩阵，因子 L 和 U 为双对角矩阵.

b. 设 $C=1$ 及 $\boldsymbol{t}_0=(10,15,15,10)^{\mathrm{T}}$，应用 A 的 LU 分解求温度分布 $\boldsymbol{t}_1$，$\boldsymbol{t}_2$，$\boldsymbol{t}_3$ 和 $\boldsymbol{t}_4$.

练习题答案

$$A=\begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix}\sim\begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & -3 & -1 & 6 \\ 0 & 6 & 2 & -7 \\ 0 & -9 & -3 & 13 \end{bmatrix}\sim\begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 10 \end{bmatrix}\sim\begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}=U$$

把每一个标志出的列除以它顶端的主元，所得的列构成 L 前三列的下半部分，这就使 L 化为 I 的变换对应于 A 化为 U 的变换，用 I_5 的后两列作为 L 的后两列，使它成为单位下三角矩阵.

$$\begin{bmatrix} 2 \\ 6 \\ 2 \\ 4 \\ -6 \end{bmatrix}\quad\begin{bmatrix} 3 \\ -3 \\ 6 \\ -9 \end{bmatrix}\quad\begin{bmatrix} 5 \\ -5 \\ 10 \end{bmatrix}$$

$$\div 2\quad\div 3\quad\div 5$$

$$\downarrow\quad\downarrow\quad\downarrow$$

$$\begin{bmatrix} 1 & & & \\ 3 & 1 & & \\ 1 & -1 & 1 & \cdots \\ 2 & 2 & -1 & \\ -3 & -3 & 2 & \end{bmatrix},\ L=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 \\ 2 & 2 & -1 & 1 & 0 \\ -3 & -3 & 2 & 0 & 1 \end{bmatrix}$$

㊀ 参见 Biswa N. Datta, *Numerical Linear Algebra and Applications* (Pacific Grove, CA: Brooks/Cole, 1994), pp. 200-201.

2.6 列昂惕夫投入产出模型

在 Wassily Leontief 获得诺贝尔奖的工作中，线性代数起着重要的作用. 如第 1 章开始所提到的，本节所叙述的经济模型是现在世界各国广泛使用模型的基础.

设某国的经济体系分为 n 个部门，这些部门生产商品和服务. 设 $\boldsymbol{x}$ 为 $\mathbb{R}^n$ 中**产出向量**，它列出了每一部门一年中的产出. 同时，设经济体系的另一部分（称为开放部门）不生产产品或服务，仅仅消费商品或服务，$\boldsymbol{d}$ 为**最终需求向量**，它列出经济体系中的各种非生产部门所需求的商品或服务. 此向量代表消费者需求、政府消费、超额生产、出口或其他外部需求.

由于各部门生产商品以满足消费者需求，生产者本身创造了**中间需求**，需要这些产品作为生产部门的投入，部门之间的关系是很复杂的，而生产和最后需求之间的联系也还不清楚. 列昂惕夫思考是否存在某一生产水平 $\boldsymbol{x}$ 恰好的满足这一生产水平的总需求（ $\boldsymbol{x}$ 称为供给），那么

$$\{\text{总产出 } \boldsymbol{x}\}=\{\text{中间需求}\}+\{\text{最终需求 } \boldsymbol{d}\} \tag{1}$$

列昂惕夫的投入产出模型的基本假设是，对每个部门，有一个单位消费向量，它列出了该部门的单位产出所需的投入. 所有的投入与产出都以百万美元作为单位，而不用具体的单位如吨等（假设商品和服务的价格为常数）.

作为一个简单的例子，设经济体系由三个部门组成——制造业、农业和服务业. 单位消费向量 $\boldsymbol{c}_1,\boldsymbol{c}_2,\boldsymbol{c}_3$ 如表 2-1 所示.

表 2-1 每单位产出消费的投入

购买自	制造业	农　业	服务业
制造业	0.50	0.40	0.20
农　业	0.20	0.30	0.10
服务业	0.10	0.10	0.30
	↑ $\boldsymbol{c}_1$	↑ $\boldsymbol{c}_2$	↑ $\boldsymbol{c}_3$

例 1 如果制造业决定生产 100 单位产品，它将消费多少？

解 计算

$$100\boldsymbol{c}_1=100\begin{bmatrix}0.50\\0.20\\0.10\end{bmatrix}=\begin{bmatrix}50\\20\\10\end{bmatrix}$$

为生产 100 单位产品，制造业需要消费制造业其他部门的 50 单位产品，20 单位农业产品，10 单位服务业产品. ■

若制造业决定生产 x_1 单位产出，则在生产的过程中消费掉的中间需求是 $x_1\boldsymbol{c}_1$，类似地，若 x_2 和 x_3 表示农业和服务业的计划产出，则 $x_2\boldsymbol{c}_2$ 和 $x_3\boldsymbol{c}_3$ 为它们的对应中间需求. 三个部门的中间需求为

$$\{\text{中间需求}\}=x_1\boldsymbol{c}_1+x_2\boldsymbol{c}_2+x_3\boldsymbol{c}_3=C\boldsymbol{x} \tag{2}$$

这里 C 是**消耗矩阵** $[\boldsymbol{c}_1 \quad \boldsymbol{c}_2 \quad \boldsymbol{c}_3]$，即

$$C=\begin{bmatrix}0.50 & 0.40 & 0.20\\ 0.20 & 0.30 & 0.10\\ 0.10 & 0.10 & 0.30\end{bmatrix} \tag{3}$$

方程（1）和（2）产生列昂惕夫模型.

列昂惕夫投入产出模型或生产方程

$$\underset{\text{总产出}}{\boldsymbol{x}} = \underset{\text{中间需求}}{C\boldsymbol{x}} + \underset{\text{最终需求}}{\boldsymbol{d}} \tag{4}$$

把 $\boldsymbol{x}$ 写成 $I\boldsymbol{x}$，应用矩阵代数，可把（4）重写为

$$\begin{aligned}I\boldsymbol{x}-C\boldsymbol{x}&=\boldsymbol{d}\\(I-C)\boldsymbol{x}&=\boldsymbol{d}\end{aligned} \tag{5}$$

例 2 考虑消耗矩阵为（3）的经济. 假设最终需求是制造业 50 单位，农业 30 单位，服务业 20 单位，求生产水平 $\boldsymbol{x}$.

解 （5）中系数矩阵为

$$I-C=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}-\begin{bmatrix}0.5 & 0.4 & 0.2\\ 0.2 & 0.3 & 0.1\\ 0.1 & 0.1 & 0.3\end{bmatrix}=\begin{bmatrix}0.5 & -0.4 & -0.2\\ -0.2 & 0.7 & -0.1\\ -0.1 & -0.1 & 0.7\end{bmatrix}$$

为解方程（5），对增广矩阵作行变换

$$\begin{bmatrix}0.5 & -0.4 & -0.2 & 50\\ -0.2 & 0.7 & -0.1 & 30\\ -0.1 & -0.1 & 0.7 & 20\end{bmatrix}\sim\begin{bmatrix}5 & -4 & -2 & 500\\ -2 & 7 & -1 & 300\\ -1 & -1 & 7 & 200\end{bmatrix}\sim\cdots\sim\begin{bmatrix}1 & 0 & 0 & 226\\ 0 & 1 & 0 & 119\\ 0 & 0 & 1 & 78\end{bmatrix}$$

最后一列四舍五入到整数，制造业需生产约 226 单位，农业 119 单位，服务业 78 单位. ■

若矩阵 $I-C$ 可逆，则我们可应用 2.2 节定理 5，用 $I-C$ 代替 A，由方程 $(I-C)\boldsymbol{x}=\boldsymbol{d}$ 得出 $\boldsymbol{x}=(I-C)^{-1}\boldsymbol{d}$. 下列定理说明，在大部分的实际情况下，$I-C$ 是可逆的，而且产出向量 $\boldsymbol{x}$ 是经济上可行的，亦即 $\boldsymbol{x}$ 中的元素是非负的.

在此定理中，**列的和**表示矩阵中某一列元素的和. 在通常情况下，某一消耗矩阵的列的和是小于 1 的，因为一个部门要生产一单位产出所需投入的总价值应该小于 1.

定理 11 设 C 为某一经济的消耗矩阵，$\boldsymbol{d}$ 为最终需求. 若 C 和 $\boldsymbol{d}$ 的元素非负，C 的每一列的和小于 1，则 $(I-C)^{-1}$ 存在，而产出向量

$$\boldsymbol{x}=(I-C)^{-1}\boldsymbol{d}$$

有非负元素，且是下列方程的唯一解

$$\boldsymbol{x}=C\boldsymbol{x}+\boldsymbol{d}$$

下列讨论说明定理成立的理由，且给出一种计算 $(I-C)^{-1}$ 的新方法.

$(I-C)^{-1}$ 的公式

假设由 $\boldsymbol{d}$ 表示的需求在年初提供给各种工业，它们制定产业水平为 $\boldsymbol{x}=\boldsymbol{d}$ 的计划，它将恰好

满足最终需求，由于这些工业准备产出为 $\boldsymbol{d}$ ，它们将提出对原料及其他投入的要求. 这就创造出对投入的需求 $C\boldsymbol{d}$.

为满足附加需求 $C\boldsymbol{d}$ ，这些工业又需要进一步的投入为 $C(C\boldsymbol{d})=C^2\boldsymbol{d}$ ，当然，它又创造出第二轮的中间需求，当要满足这些需求时，它们又创造出第三轮需求，即 $C(C^2\boldsymbol{d})=C^3\boldsymbol{d}$ ，等等.

理论上，这个过程可无限延续下去，虽然实际上这样一系列事件不可能一直发生下去. 我们可把这一假设的情形表示如表 2-2 所示.

表 2-2

	要满足的需求	为满足此需求需要的投入
最终需求	$\boldsymbol{d}$	$C\boldsymbol{d}$
中间需求		
第一轮	$C\boldsymbol{d}$	$C(C\boldsymbol{d})=C^2\boldsymbol{d}$
第二轮	$C^2\boldsymbol{d}$	$C(C^2\boldsymbol{d})=C^3\boldsymbol{d}$
第三轮	$C^3\boldsymbol{d}$	$C(C^3\boldsymbol{d})=C^4\boldsymbol{d}$
	$\vdots$	$\vdots$

为了满足所有这些需求的产出水平 $\boldsymbol{x}$ 是

$$\boldsymbol{x}=\boldsymbol{d}+C\boldsymbol{d}+C^2\boldsymbol{d}+C^3\boldsymbol{d}+\cdots=(I+C+C^2+C^3+\cdots)\,\boldsymbol{d} \tag{6}$$

为了使（6）有意义，我们使用下列代数恒等式：

$$(I-C)(I+C+C^2+\cdots+C^m)=I-C^{m+1} \tag{7}$$

可以证明，若 C 的列的和都严格小于 1，则 $I-C$ 是可逆的，当 m 趋于无穷时 C^m 趋于 0，而 $I-C^{m+1}\to I$.（这有点类似于当正数 t 小于 1 时，随着 m 增大， $t^m\to 0$.）应用（7），我们有

$$\text{当 } C \text{ 的列的和小于 1 时，} (I-C)^{-1}\approx I+C+C^2+C^3+\cdots+C^m \tag{8}$$

我们将（8）解释为当 m 充分大时，右边可以任意接近于 $(I-C)^{-1}$.

在实际的投入产出模型中，消耗矩阵的幂迅速趋于 0，故（8）实际上给出一种计算 $(I-C)^{-1}$ 的方法. 类似地，对任意 $\boldsymbol{d}$ ，向量 $C^m\boldsymbol{d}$ 迅速地趋于零向量，而（6）给出实际解 $(I-C)\boldsymbol{x}=\boldsymbol{d}$ 的方法. 若 C 和 $\boldsymbol{d}$ 中的元素是非负的，则（6）说明 $\boldsymbol{x}$ 中的元素也是非负的.

$(I-C)^{-1}$ 中元素的经济重要性

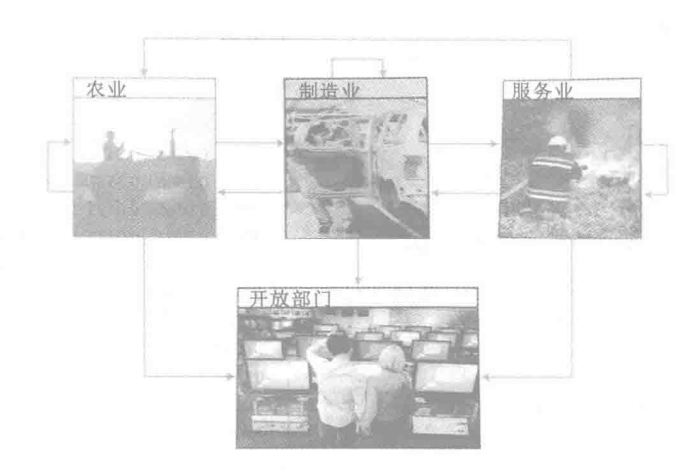
$(I-C)^{-1}$ 中的元素是有意义的，因它们可用来预计当最终需求 $\boldsymbol{d}$ 改变时，产出向量 $\boldsymbol{x}$ 如何改变. 事实上，$(I-C)^{-1}$ 的第 j 列表示当第 j 个部门的最终需求增加 1 单位时，各部门需要增加产出的数量. 见习题 8.

数值计算的注解 在任何应用问题中（不仅是经济学）方程 $A\boldsymbol{x}=\boldsymbol{b}$ 总可以写成 $(I-C)\boldsymbol{x}=\boldsymbol{d}$ 的形式，其中 $C=I-A$. 若方程组很大而且稀疏（大部分元素为 0），可能 C 的各列元素绝对值之和小于 1，这时 $C^m\to 0$，若 C^m 趋向于零足够迅速，（6）和（8）可以用来作为解方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的实际方法，也可用来求 A^{-1} .

练习题

设某一经济有两个部门，商品和服务部门. 商品部门的单位产出需要 0.2 单位商品和 0.5 单位服务的投入，服务部门的单位产出需要 0.4 单位商品和 0.3 单位服务的投入. 最终需求是 20 单位商品和 30 单位服务，列出列昂惕夫投入产出模型的方程.

习题 2.6



习题 1~4 讨论一个经济体系，它分为制造业、农业和服务业三个部门. 见上图. 制造业每单位产出需要 0.10 单位制造业产品，0.30 单位农业产品和 0.30 单位服务产品投入. 每单位农业产出需要 0.20 单位它自己的产出，0.60 单位制造业产出，0.10 单位服务产出，服务业的每单位产出消耗 0.10 单位服务，0.60 单位制造业产品，但不消耗农业产出.

1. 构造此经济的消耗矩阵，若农业要生产 100 单位产出，产生的中间需求是什么？
2. 为了满足最终需求为 20 单位农业产品（对其他部门无最终需求），总的产出水平应为多少（不要计算逆矩阵）.
3. 为了满足最终需求为 20 单位制造业产品（对其他部门无最终需求），总的产出水平应为多少（不要计算逆矩阵）.
4. 为了满足最终需求为 20 单位制造业产品，20 单位农业产品，0 单位服务，总的产出水平应为多少.
5. 考虑生产模型 $\boldsymbol{x}=C\boldsymbol{x}+\boldsymbol{d}$，该经济体系有两个部门，其中
$$C=\begin{bmatrix}0.0 & 0.5\\0.6 & 0.2\end{bmatrix},\quad \boldsymbol{d}=\begin{bmatrix}50\\30\end{bmatrix}$$
应用逆矩阵来确定最终需求.
6. 重复习题 5，取 $C=\begin{bmatrix}0.2 & 0.5\\0.6 & 0.1\end{bmatrix}$，$\boldsymbol{d}=\begin{bmatrix}16\\12\end{bmatrix}$.
7. 设 C 和 $\boldsymbol{d}$ 如题 5.
   a. 为满足最终需求为部门 1 的 1 单位产品，产出水平应为多少？
   b. 用逆矩阵求出最终需求为 $\begin{bmatrix}51\\30\end{bmatrix}$ 时的总产出水平.
   c. 应用事实 $\begin{bmatrix}51\\30\end{bmatrix}=\begin{bmatrix}50\\30\end{bmatrix}+\begin{bmatrix}1\\0\end{bmatrix}$ 说明（a）和（b）的答案以及题 5 的答案有何关系.
8. 设 C 为 $n\times n$ 消耗矩阵，它的列的和小于 1，设 $\boldsymbol{x}$ 为满足最终需求 $\boldsymbol{d}$ 的产出向量，$\Delta\boldsymbol{x}$ 为满足不同的最终需求 $\Delta\boldsymbol{d}$ 的产出向量.
   a. 证明：若最终需求由 $\boldsymbol{d}$ 改变为 $\boldsymbol{d}+\Delta\boldsymbol{d}$，则新

的产出水平必须为 $x+\Delta x$，于是 Δx 给出需求改变 Δd 时产出改变的量.

b. 设 Δd 为 $\mathbb{R}^n$ 中第 1 个元素为 1 其他元素为 0 的向量，说明为什么对应的产出 Δx 是 $(I-C)^{-1}$ 的第 1 列，这证明 $(I-C)^{-1}$ 的第 1 列给出当第 1 部门的最终需求增加 1 单位时，其他部门需要增加的产出.

9. 设某一经济有 3 个部门，解它的列昂惕夫方程，给出

$$C=\begin{bmatrix}0.2 & 0.2 & 0.0\\ 0.3 & 0.1 & 0.3\\ 0.1 & 0.0 & 0.2\end{bmatrix},\quad d=\begin{bmatrix}40\\60\\80\end{bmatrix}$$

10. 美国经济在 1972 年的消耗矩阵 C 具有下列性质，矩阵 $(I-C)^{-1}$ 的每个元素都是正的. 这说明仅增加某一部门的产出的需求对其他部门的影响是什么.[㊀]

11. 列昂惕夫产出方程 $x=Cx+d$ 通常与对偶的**价格方程**

$$p=C^{\mathrm{T}}p+v$$

联系在一起，其中 p 为**价格向量**，它的元素列出各部门产出的单位价格，v 是**增值向量**，它的元素是每单位产出附加的价值（增值包括工资、利润、折旧等）. 经济中重要的事实是国内生产总值（GDP）可用两种方式表示：

$$\{国内生产总值\}=p^{\mathrm{T}}d=v^{\mathrm{T}}x$$

证明第二个等式.（提示：用两种方式计算 $p^{\mathrm{T}}x$.）

12. 设 C 为消耗矩阵，当 $m\to\infty$ 时 $C^m\to 0$，对 $m=1,2,\cdots$，令 $D_m=I+C+\cdots+C^m$，求出把 D_m 和 D_{m+1} 联系的差分方程，得出由（8）计算 $(I-C)^{-1}$ 的迭代算法.

13. [M]下列的消耗矩阵 C 是基于 1958 年美国经济的投入产出数据，把 81 个部门合并成 7 个大的部门：（1）非金属家用及个人产品，（2）最终金属产品（如汽车），（3）基础金属产品及矿业，（4）基础非金属产品与农业，（5）能源，（6）服务业，（7）娱乐与其他产品.[㊁] 求出满足最终需求 d 的产出水平（单位：百万美元）.

$$C=\begin{bmatrix}0.1588 & 0.0064 & 0.0025 & 0.0304 & 0.0014 & 0.0083 & 0.1594\\ 0.0057 & 0.2645 & 0.0436 & 0.0099 & 0.0083 & 0.0201 & 0.3413\\ 0.0264 & 0.1506 & 0.3557 & 0.0139 & 0.0142 & 0.0070 & 0.0236\\ 0.3299 & 0.0565 & 0.0495 & 0.3636 & 0.0204 & 0.0483 & 0.0649\\ 0.0089 & 0.0081 & 0.0333 & 0.0295 & 0.3412 & 0.0237 & 0.0020\\ 0.1190 & 0.0901 & 0.0996 & 0.1260 & 0.1722 & 0.2368 & 0.3369\\ 0.0063 & 0.0126 & 0.0196 & 0.0098 & 0.0064 & 0.0132 & 0.0012\end{bmatrix},$$

$$d=\begin{bmatrix}74\,000\\56\,000\\10\,500\\25\,000\\17\,500\\196\,000\\5\,000\end{bmatrix}$$

14. [M]习题 13 中的需求向量是 1958 年数据. 但列昂惕夫在讨论中应用了一个接近 1964 年的数据：

$$d=(99\,640,75\,548,14\,444,33\,501,23\,527,263\,985,6\,526)$$

求出满足此需求的产出水平.

15. [M]用方程（6）解 13 题，设 $x^{(0)}=d$，对 $k=1,2,\cdots$，计算 $x^{(k)}=d+Cx^{(k-1)}$，需要多少步才能得到有 4 位有效数字的答案？

练习题答案

已给下列数据：

购买自	每单位产出所需投入		外部需求
	商品	服务	
商品	0.2	0.4	20
服务	0.5	0.3	30

㊀ Wassily W.Leontief, "The World Economy of the Year 2000. "*Scientific American,*September 1980, pp. 206-231.

㊁ Wassily W.Leontief, "The Structure of the U.S.Economy,"*Scientific American*,April,1965, pp. 30-32.

列昂惕夫投入产出模型为 $\boldsymbol{x}=C\boldsymbol{x}+\boldsymbol{d}$，其中

$$C=\begin{bmatrix}0.2 & 0.4\\0.5 & 0.3\end{bmatrix},\quad \boldsymbol{d}=\begin{bmatrix}20\\30\end{bmatrix}$$

2.7 计算机图形学中的应用

计算机图形是在计算机屏幕上显示或活动的图像. 计算机图形学的应用广泛，发展迅速. 例如，计算机辅助设计（CAD）是许多工程技术的组成部分之一，比如本章介绍中提到的飞机设计过程. 娱乐行业对计算图形学做了最引人入胜的应用——从黑客帝国的特技效果到 PlayStation 2 电脑娱乐系统和 Xbox 游乐器游戏.

绝大多数工业或商业的交互计算机软件在屏幕上应用计算机图形显示以及其他功能，如数据的图形显示、桌面编辑以及商业或教育用的幻灯片等，因此，任何学习计算机语言的学生至少要学会如何应用二维（2D）图形.

本节考虑用来操纵和显示图形图像的一些基本的数学方法，例如飞机的线形轮廓模型. 这样的一个图像（图片）是由一系列的点和曲线组成，以及如何填充由直线和曲线所围成的封闭区域. 通常，曲线用短的直线段逼近，而图形用一系列的点来定义.

在最简单的二维图形符号中，字母用于在屏幕上做标记. 某些字母作为线框对象存储，其他有弯曲部分的字母还要将曲线的数学公式也存储进去.

例 1 图 2-15 中的大写字母 N 由 8 个点或顶点确定. 这些点的坐标可存储在一个数据矩阵 D 中.

$$\begin{matrix}\text{顶点} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8\\ x\text{ 坐标} \\ y\text{ 坐标}\end{matrix}\quad \begin{bmatrix}0 & 0.5 & 0.5 & 6 & 6 & 5.5 & 5.5 & 0\\0 & 0 & 6.42 & 0 & 8 & 8 & 1.58 & 8\end{bmatrix}=D$$

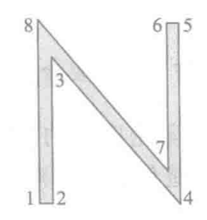


图 2-15 常规的 N

除 D 以外，还要说明哪些顶点用线相连，但我们省略这些细节. ■

图形对象使用一组直线线段描述的主要原因是，计算机图形学中标准变换把线段映射成为线段.（例如，见 1.8 节习题 26.）当描述这些对象的顶点被变换以后，它们的像可以用适当的直线连接起来得到原来对象的完整图像.

例 2 给定 $A=\begin{bmatrix}1 & 0.25\\0 & 1\end{bmatrix}$ 描述剪切变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 对例 1 中字母 N 的作用.

解 由矩阵乘法的定义，乘积 AD 的各列给出字母 N 各顶点的像.

$$AD = \begin{bmatrix} \overset{1}{0} & \overset{2}{0.5} & \overset{3}{2.105} & \overset{4}{6} & \overset{5}{8} & \overset{6}{7.5} & \overset{7}{5.895} & \overset{8}{2} \\ 0 & 0 & 6.420 & 0 & 8 & 8 & 1.580 & 8 \end{bmatrix}$$

变换过的顶点画在图 2-16，同时还画上相应于原来图形中连线的线段. ■

图 2-16 中斜体的 N 看来有些太宽，为此，我们可以用倍乘变换使它变窄.

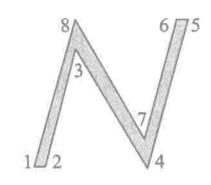


图 2-16 斜体的 N

例 3 先作如例 2 的剪切变换，然后再把 x 坐标乘以一个因子 0.75，求此复合变换的矩阵.

解 把每个点的 x 坐标乘以 0.75 的矩阵为

$$S = \begin{bmatrix} 0.75 & 0 \\ 0 & 1 \end{bmatrix}$$

所以复合变换的矩阵是

$$SA = \begin{bmatrix} 0.75 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0.25 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.75 & 0.1875 \\ 0 & 1 \end{bmatrix}$$

复合变换的结果如图 2-17 所示.

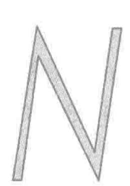
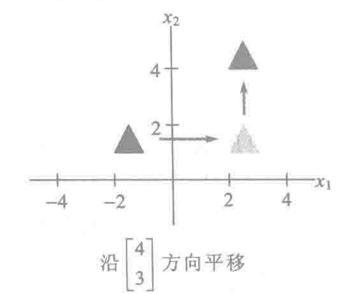


图 2-17 N 的复合变换 ■

计算机图形学中的数学是与矩阵乘法紧密联系的. 但是，屏幕上的物体的平移并不直接对应于矩阵乘法，因为平移并非线性变换. 避免这一困难的标准办法是引入所谓齐次坐标.

齐次坐标

$\mathbb{R}^2$ 中每个点 (x, y) 可以对应于 $\mathbb{R}^3$ 中的 $(x, y, 1)$. 它们位于 xy 平面上方 1 单位的平面上. 我们称 (x, y) 有齐次坐标 $(x, y, 1)$，例如，点 $(0,0)$ 的齐次坐标为 $(0,0,1)$. 点的齐次坐标不能相加，也不能乘以数，但它们可以乘以 3×3 矩阵来做变换.



沿 $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$ 方向平移

例 4 形如 $(x, y) \mapsto (x+h, y+k)$ 的平移可以用齐次坐标写成 $(x, y, 1) \mapsto (x+h, y+k, 1)$，这个变换可用矩阵乘法来实现：

$$\begin{bmatrix}1&0&h\\0&1&k\\0&0&1\end{bmatrix}\begin{bmatrix}x\\y\\1\end{bmatrix}=\begin{bmatrix}x+h\\y+k\\1\end{bmatrix}$$ ■

例 5 $\mathbb{R}^2$ 中任意线性变换也可用通过齐次坐标乘以分块矩阵 $\begin{bmatrix}A&0\\0&1\end{bmatrix}$ 实现，其中 A 是 2×2 矩阵，典型的例子是：

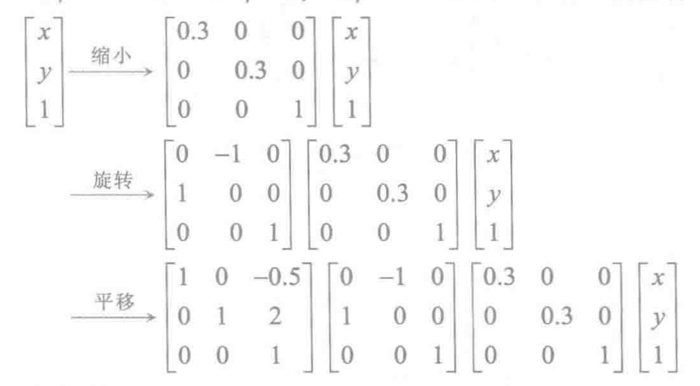
$$\underset{\text{绕原点逆时针旋转角度}\varphi}{\begin{bmatrix}\cos\varphi&-\sin\varphi&0\\\sin\varphi&\cos\varphi&0\\0&0&1\end{bmatrix}},\quad \underset{\text{关于 }y=x\text{ 的对称}}{\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}},\quad \underset{x\text{ 乘以 }s,\ y\text{ 乘以 }t}{\begin{bmatrix}s&0&0\\0&t&0\\0&0&1\end{bmatrix}}$$ ■

复合变换

图形在计算机屏幕上的移动通常需要两个或多个基本变换. 这些变换的复合相应于在使用齐次坐标时进行矩阵相乘.

例 6 求出 3×3 矩阵，对应于先乘以 0.3 的倍乘变换，然后旋转 90°，最后对图形的每个点的坐标加上 (–0.5, 2) 做平移. 见图 2-18.

解 当 $\varphi=\pi/2$ 时，$\sin\varphi=1,\cos\varphi=0$，由例 4 和 5，我们有



$$\begin{bmatrix}x\\y\\1\end{bmatrix}\xrightarrow{\text{缩小}}\begin{bmatrix}0.3&0&0\\0&0.3&0\\0&0&1\end{bmatrix}\begin{bmatrix}x\\y\\1\end{bmatrix}$$

$$\xrightarrow{\text{旋转}}\begin{bmatrix}0&-1&0\\1&0&0\\0&0&1\end{bmatrix}\begin{bmatrix}0.3&0&0\\0&0.3&0\\0&0&1\end{bmatrix}\begin{bmatrix}x\\y\\1\end{bmatrix}$$

$$\xrightarrow{\text{平移}}\begin{bmatrix}1&0&-0.5\\0&1&2\\0&0&1\end{bmatrix}\begin{bmatrix}0&-1&0\\1&0&0\\0&0&1\end{bmatrix}\begin{bmatrix}0.3&0&0\\0&0.3&0\\0&0&1\end{bmatrix}\begin{bmatrix}x\\y\\1\end{bmatrix}$$

图 2-18

所以复合变换的矩阵为

$$\begin{bmatrix}1&0&-0.5\\0&1&2\\0&0&1\end{bmatrix}\begin{bmatrix}0&-1&0\\1&0&0\\0&0&1\end{bmatrix}\begin{bmatrix}0.3&0&0\\0&0.3&0\\0&0&1\end{bmatrix}=\begin{bmatrix}0&-1&-0.5\\1&0&2\\0&0&1\end{bmatrix}\begin{bmatrix}0.3&0&0\\0&0.3&0\\0&0&1\end{bmatrix}=\begin{bmatrix}0&-0.3&-0.5\\0.3&0&2\\0&0&1\end{bmatrix}$$ ■

三维计算机图形学

计算机图形学的最新和最激动人心的应用是分子模型. 对三维图形学，生物学家可观察到模拟的蛋白质分子，并用以研究药物分子. 生物学家可以旋转和平移一种实验药物的分子使它们附着于蛋白质分子. 研究潜在的化学反应的能力对现代药物和癌症研究是很重要的. 事实

上，药物设计的进展依赖于计算机图形学构造有真实感的分子和它们的交互作用的仿真.[⊖]

现在的分子模型的研究集中于虚拟现实，即一种环境，研究者在其中可以看到并且感觉药物分子渗入蛋白质分子，在图 2-19 中，这样的有实体感的反馈是由能够将力量可视化的远程控制器提供的. 另一种虚拟现实的设计包括用一个头盔和手套来感觉头、手和手指的运动. 头盔包含两个细小的计算机屏幕，每个眼睛一个，将此可视环境变得更为真实，这是对工程师，科学家与数学家更大的挑战. 我们这里考虑的数学打开了这方面的研究的大门.

图 2-19 虚拟现实中的分子模型（北卡罗来纳大学 Chapel Hill 分校计算机科学系，照片由 Bo Strain 提供）

齐次三维坐标

类似于二维情形，我们称 $(x,y,z,1)$ 是 $\mathbb{R}^3$ 中点 (x,y,z) 的齐次坐标. 一般地，若 $H\neq 0$，则 (X,Y,Z,H) 是 (x,y,z) 的齐次坐标，且

$$x=\frac{X}{H},\quad y=\frac{Y}{H},\quad z=\frac{Z}{H} \tag{1}$$

$(x,y,z,1)$ 的每一个非零的标量乘法得到一组 (x,y,z) 的齐次坐标. 例如 $(10,-6,14,2)$ 和 $(-15,9,-21,-3)$ 都是 $(5,-3,7)$ 的齐次坐标.

下列说明分子模型中，把一个药物分子移入蛋白质分子的变换.

例 7 给出下列变换的 4×4 矩阵.

a. 绕 y 轴旋转 30°（习惯上，正角是从旋转轴（本例中是 y 轴）的正半轴向原点看过去的逆时针方向的角）.

b. 沿向量 $\boldsymbol{p}=(-6,4,5)$ 的方向平移.

解 a. 首先构造 3×3 矩阵表示旋转. 向量 $\boldsymbol{e}_1$ 向负 z 轴旋转到$(\cos 30°,0,-\sin 30°)=(\sqrt{3}/2,0,-0.5)$，

⊖ Robert Pool, "Computing in Science", *Science* 256, 3 April 1992, p.45.

向量 $\boldsymbol{e}_2$ 不变，$\boldsymbol{e}_3$ 向正 x 轴旋转到 $(\sin 30°, 0, \cos 30°) = (0.5, 0, \sqrt{3}/2)$．见图 2-20．由 1.9 节，这个旋转变换的标准矩阵为

$$A = \begin{bmatrix} \sqrt{3}/2 & 0 & 0.5 \\ 0 & 1 & 0 \\ -0.5 & 0 & \sqrt{3}/2 \end{bmatrix}$$

所以齐次坐标的旋转矩阵为

$$A = \begin{bmatrix} \sqrt{3}/2 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 0 \\ -0.5 & 0 & \sqrt{3}/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

b. 我们希望 $(x, y, z, 1)$ 映射到 $(x-6, y+4, z+5, 1)$，所求矩阵为

$$\begin{bmatrix} 1 & 0 & 0 & -6 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

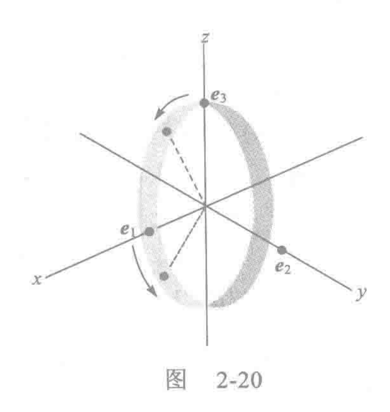


图 2-20 ■

透视投影

三维物体在二维计算机屏幕上的表示方法是把它投影在一个可视平面上.（我们忽略其他重要步骤，例如选择可视平面上显示在屏幕上的部分.）为简单起见，设 xy 平面表示计算机屏幕，假设某一观察者的眼睛向正 z 轴看去，眼睛的位置是 $(0,0,d)$，透视投影把每个点 (x, y, z) 映射为点 $(x^*, y^*, 0)$，使这两点与观察者的眼睛位置（称为透视中心）在一条直线上，见图 2-21（a）.

xz 平面上的三角形（在图 2-21a 中）画在 b 中，由相似三角形知

$$\frac{x^*}{d} = \frac{x}{d-z} \quad , \quad x^* = \frac{dx}{d-z} = \frac{x}{1 - z/d}$$

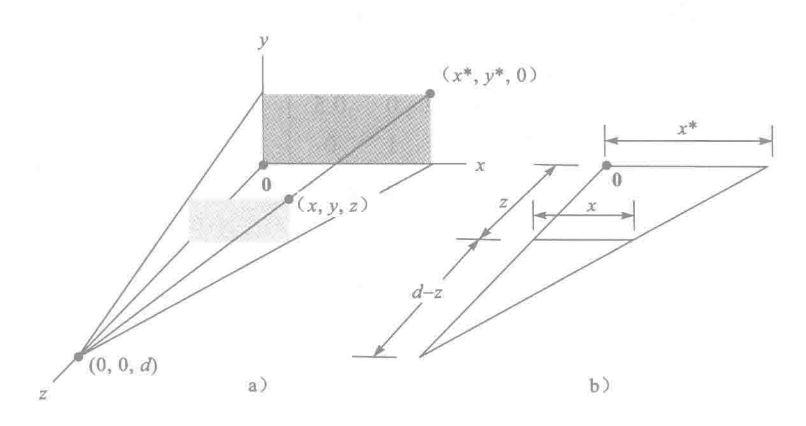


图 2-21 由 (x,y,z) 到 $(x^*,y^*,0)$ 的透视投影

类似地，

$$y^* = \frac{y}{1-z/d}$$

使用齐次坐标，可用矩阵表示透视投影，记此矩阵为 P，$(x,y,z,1)$ 映射为

$$\left(\frac{x}{1-z/d},\frac{y}{1-z/d},0,1\right)$$

把这个向量乘以 $1-z/d$，可用 $(x,y,0,1-z/d)$ 作为齐次坐标的像，现在容易求出 P. 事实上

$$P\begin{bmatrix}x\\y\\z\\1\end{bmatrix}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&-1/d&1\end{bmatrix}\begin{bmatrix}x\\y\\z\\1\end{bmatrix}=\begin{bmatrix}x\\y\\0\\1-z/d\end{bmatrix}$$

例 8 设 S 是顶点为(3, 1, 5)，(5, 1, 5)，(5, 0, 5)，(3, 0, 5)，(3, 1, 4)，(5, 1, 4)及(3, 0, 4)的长方体，如图 2-22 所示，求 S 在透视中心为(0, 0, 10)的透视投影下的像.

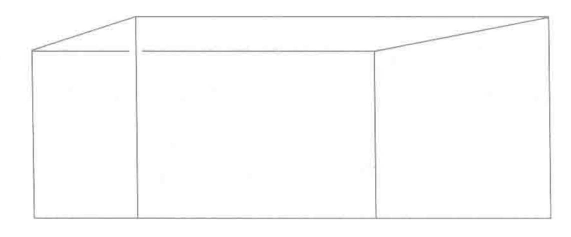


图 2-22 透视投影下的 S

解 设 P 为投影矩阵，D 为用齐次坐标的 S 的数据矩阵，则 S 的像的数据矩阵为

$$
PD = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1/10 & 1 \end{bmatrix} \overset{\text{顶点}}{\overset{1\ \ 2\ \ 3\ \ 4\ \ 5\ \ 6\ \ 7\ \ 8}{\begin{bmatrix} 3 & 5 & 5 & 3 & 3 & 5 & 5 & 3 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 5 & 4 & 4 & 4 & 4 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}}} = \begin{bmatrix} 3 & 5 & 5 & 3 & 3 & 5 & 5 & 3 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0.5 & 0.5 & 0.6 & 0.6 & 0.6 & 0.6 \end{bmatrix}
$$

为得到 $\mathbb{R}^3$ 坐标，使用例 7 前面的（1）式，把每一列的前 3 个元素除以第 4 行的对应元素，得到

$$
\overset{\text{顶点}}{\overset{1\ \ 2\ \ 3\ \ 4\ \ 5\ \ 6\ \ 7\ \ 8}{\begin{bmatrix} 6 & 10 & 10 & 6 & 5 & 8.3 & 8.3 & 5 \\ 2 & 2 & 0 & 0 & 1.7 & 1.7 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}}}
$$

■

本教材的网页上有计算机图形学的有趣应用，包括对透视投影的进一步讨论. 网上的一个计算机项目与简单的动画有关.

数值计算的注解 三维物体的连续移动需要计算大量的 4×4 矩阵，特别地，当曲面光滑时，用来使它更有实体感，并有适当的光线，图形工作站有 4×4 矩阵运算及图形算法嵌入于芯片和电路中，这样的工作站可以每秒做数十亿次矩阵乘法以实现三维游戏程序中有真实感的颜色的变化.[⊖]

进一步阅读

James D. Foley, Andries van Dam, Steven K. Feiner, and John F. Hughes, *Computer Graphics: Principles and Practice*, 3rd ed. (Boston, MA: Addison-Wesley, 2002), Chapters 5 and 6.

练习题

图形绕 $\mathbb{R}^2$ 中一点 $\boldsymbol{p}$ 的旋转是这样实现的：首先把图形平移 $-\boldsymbol{p}$，然后绕原点旋转，最后平移回去 $\boldsymbol{p}$. 见图 2-23，使用齐次坐标构造绕点(−2, 6)旋转−30° 的 3×3 矩阵.

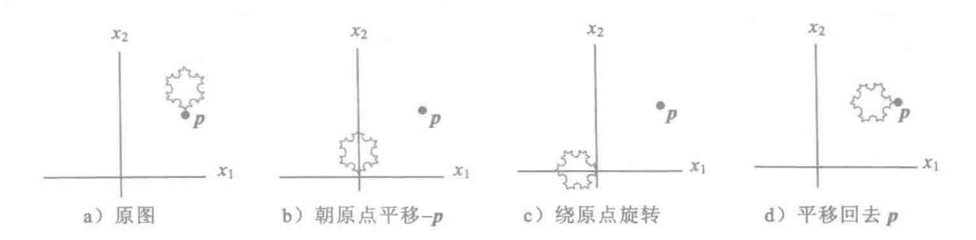


图 2-23 图形绕点 $\boldsymbol{p}$ 旋转

⊖ 见 Jan Ozer, "High-Performance Graphics Boards", *PC Magazine* 19, 1 September 2000, pp. 187-200. 也见"The Ultimate Upgrade Guide: Moving On Up", *PC Magazine* 21, 29 January 2002, pp. 82-91.

习题 2.7

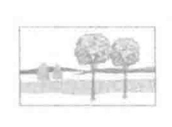
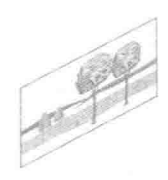
1. 哪一个 3×3 矩阵对 $\mathbb{R}^2$ 齐次坐标的作用与例 2 中的剪切变换矩阵 A 相同？
2. 用矩阵乘法求出由数据矩阵 $D=\begin{bmatrix}4 & 2 & 5\\0 & 2 & 3\end{bmatrix}$ 决定的三角形在关于 y 轴的对称变换下的像. 画出原来的三角形和它的像.

习题 3~8 中，求出产生所述复合二维变换的 3×3 矩阵，用齐次坐标.

3. 先沿(2, 1)方向平移，然后绕原点旋转 90°.
4. 先沿(−1, 4)方向平移，然后把 x 坐标乘 1/2，y 坐标乘 3/2.
5. 先关于 x 轴对称，然后绕原点旋转 45°.
6. 先绕原点旋转 45°，再关于 x 轴对称.
7. 绕点(6, 8)旋转 60°.
8. 绕点(3, 7)旋转 45°.
9. 2×200 的数据矩阵 D 包含 100 个点的坐标. 要把这些点用 2×2 的任意矩阵 A 和 B 变换，需要多少次乘法？考虑两种方法 $A(BD)$ 和 $(AB)D$，讨论你的结果对计算机图形学计算的含意.
10. 考虑下列二维变换：D 是拉伸变换（x 轴与 y 轴同时乘一个数），R 是旋转变换，T 是平移变换，D 是否与 R 可交换？即是否对 $\mathbb{R}^2$ 中一切 $\boldsymbol{x}$ 有 $D(R(\boldsymbol{x}))=R(D(\boldsymbol{x}))$，$D$ 是否与 T 可交换？R 是否与 T 可交换？
11. 计算机屏幕上的一个旋转变化有时可用两个剪切-拉伸变换的复合来实现. 这样可以加快计算从而确定某一图像如何以屏幕像素显示在计算机屏幕上.（屏幕由按行和列排列的小的点组成，称为像素.）第一个变换 A_1 垂直地剪切然后压缩每列的像素；第二个变换 A_2 水平剪切然后拉伸每行的像素. 设

$$A_1=\begin{bmatrix}1 & 0 & 0\\ \sin\varphi & \cos\varphi & 0\\ 0 & 0 & 1\end{bmatrix},\quad A_2=\begin{bmatrix}\sec\varphi & -\tan\varphi & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

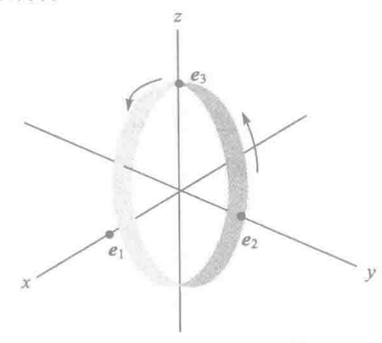
证明：这两个变换的复合是 $\mathbb{R}^2$ 中的旋转变换.

12. $\mathbb{R}^2$ 中的旋转变换通常需要 4 次乘法，计算下面的乘积，证明：一个旋转矩阵可以分解为三个剪切变换的乘积（每一个剪切变换只需作一次乘法）.

$$\begin{bmatrix}1 & -\tan\varphi/2 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ \sin\varphi & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & -\tan\varphi/2 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

13. 通常二维图形的齐次坐标变换涉及形如 $\begin{bmatrix}A & \boldsymbol{p}\\ \boldsymbol{0}^{\mathrm{T}} & 1\end{bmatrix}$ 的 3×3 矩阵. A 为 2×2 矩阵，$\boldsymbol{p}$ 是 $\mathbb{R}^2$ 中的向量. 证明这样一个变换等同于在 $\mathbb{R}^2$ 中的一个线性变换之后再作一个平移.（提示：求出包含分块矩阵的一个适当的矩阵分解.）
14. 证明：题 7 中的变换等价于绕原点的一个旋转之后再沿方向 $\boldsymbol{p}$ 平移，求出向量 $\boldsymbol{p}$.
15. $\mathbb{R}^3$ 中什么向量有齐次坐标 $(1/2,-1/4,1/8,1/24)$？
16. $(1,-2,3,4)$ 与 $(10,-20,-30,40)$ 是 $\mathbb{R}^3$ 中同一个点的齐次坐标吗？为什么？
17. 给出一个 4×4 矩阵，它将 $\mathbb{R}^3$ 中的点绕 x 轴逆时针旋转 60°.



18. 给出 4×4 矩阵，它将 $\mathbb{R}^3$ 中的点绕 z 轴旋转

$-30°$，然后沿 $\boldsymbol{p}=(5,-2,1)$ 方向平移.

19. 设 S 是顶点为 $(4.2,1.2,4)$，$(6,4,2)$，$(2,2,6)$ 的三角形，求出透视中心在 $(0,0,10)$ 处时 S 的透视投影的像.
20. 设 S 是顶点为 $(7,3,-5)$，$(12,8,2)$，$(1,2,1)$ 的三角形，求出透视中心在 $(0,0,10)$ 处时 S 的透视投影的像.

习题 21 和习题 22 考虑在计算机图形学中显示颜色的设置方法. 计算机屏幕上的颜色是由 3 个数 (R,G,B) 确定的，它列出电子枪射击到屏幕上的红、绿、蓝荧光点的能量的大小.（第 4 个数说明颜色的亮度或强度.）

21. [M]观察者在屏幕上看到的实际颜色是被屏幕上荧光点的数量与类型决定的. 每个计算机显示器制造商必须在数据 (R,G,B) 和使用三原色 X,Y,Z 的国际 CIE 颜色标准之间转换. 一种对持续时间短暂的荧光的转换是

$$\begin{bmatrix}0.61 & 0.29 & 0.150\\0.35 & 0.59 & 0.063\\0.04 & 0.12 & 0.787\end{bmatrix}\begin{bmatrix}R\\G\\B\end{bmatrix}=\begin{bmatrix}X\\Y\\Z\end{bmatrix}$$

计算机程序将传送颜色信息到屏幕上，使用标准 CIE 数据 (X,Y,Z). 求出把这些数据转换为屏幕电子枪所需的 (R,G,B) 数据的方程.

22. [M]商业电视的信号广播用向量 (Y,I,Q) 描述颜色. 若屏幕是黑白的，则只使用 Y 坐标.（它给出比用 CIE 颜色数据更好的单色图像.）YIQ 与"标准"的 RGB 颜色的对应关系是

$$\begin{bmatrix}Y\\I\\Q\end{bmatrix}=\begin{bmatrix}0.299 & 0.587 & 0.114\\0.596 & -0.275 & -0.321\\0.212 & -0.528 & 0.311\end{bmatrix}\begin{bmatrix}R\\G\\B\end{bmatrix}$$

（显示器制造商将改变矩阵元素使之适合它的 RGB 屏幕.）求出由电视台发送的 YIQ 数据转换为电视机屏幕所需的 RGB 数据的方程.

练习题答案

从右向左将这三个变换对应的矩阵组合在一起. 使用 $\boldsymbol{p}=(-2,6),\cos(-30°)=\sqrt{3}/2,\sin(-30°)=-0.5$，我们有

$$\underset{\text{平移}\ \boldsymbol{p}}{\begin{bmatrix}1 & 0 & -2\\0 & 1 & 6\\0 & 0 & 1\end{bmatrix}}\underset{\text{绕原点旋转}}{\begin{bmatrix}\sqrt{3}/2 & 1/2 & 0\\-1/2 & \sqrt{3}/2 & 0\\0 & 0 & 1\end{bmatrix}}\underset{\text{平移}-\boldsymbol{p}}{\begin{bmatrix}1 & 0 & 2\\0 & 1 & -6\\0 & 0 & 1\end{bmatrix}}=\begin{bmatrix}\sqrt{3}/2 & 1/2 & \sqrt{3}-5\\-1/2 & \sqrt{3}/2 & -3\sqrt{3}+5\\0 & 0 & 1\end{bmatrix}$$

2.8 $\mathbb{R}^n$ 的子空间

本节讨论 $\mathbb{R}^n$ 中重要的向量子集，称为子空间. 通常子空间与某个矩阵 A 有关，它们提供了关于方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的有用信息. 本节的概念和术语将在本书以下部分经常出现.[㊀]

定义 $\mathbb{R}^n$ 中的一个子空间是 $\mathbb{R}^n$ 中的集合 H，具有以下三个性质：

a. 零向量属于 H.

b. 对 H 中任意的向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$，$\boldsymbol{u}+\boldsymbol{v}$ 属于 H.

㊀ 2.8 节和 2.9 节放在这里，使读者可以跳过以下两章的大部分内容直接进入第 5 章. 如果读者计划先读第 3 章和第 4 章，再读第 5 章，则可以跳过这两节.

c. 对 H 中任意向量 $\boldsymbol{u}$ 和数 c，$c\boldsymbol{u}$ 属于 H .

换句话说，子空间对加法和标量乘法运算是封闭的. 你将在以下例子看到，第 1 章中所讨论的向量集合大部分是子空间. 例如，通过原点的一个平面是一种很典型的子空间，可以看作例 1 中子空间的一种实例. 见图 2-24.

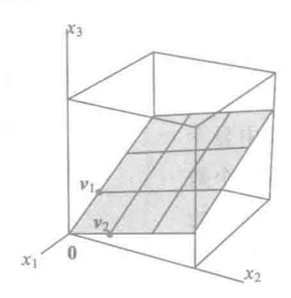


图 2-24 Span$\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ 是通过原点的平面

例 1 若 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 是 $\mathbb{R}^n$ 中的向量，$H = \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$，则 H 是 $\mathbb{R}^n$ 的子空间. 为证明这一点，注意零向量属于 H（因 $0\boldsymbol{v}_1 + 0\boldsymbol{v}_2$ 是 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的线性组合），现取 H 中任意两个向量，比如说

$$\boldsymbol{u} = s_1\boldsymbol{v}_1 + s_2\boldsymbol{v}_2 \text{ ,} \boldsymbol{v} = t_1\boldsymbol{v}_1 + t_2\boldsymbol{v}_2 \text{ ,}$$

那么

$$\boldsymbol{u} + \boldsymbol{v} = (s_1 + t_1)\boldsymbol{v}_1 + (s_2 + t_2)\boldsymbol{v}_2$$

这证明了 $\boldsymbol{u}+\boldsymbol{v}$ 是 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的线性组合，因此属于 H，同样，对任意数 c，向量 $c\boldsymbol{u}$ 属于 H，因为 $c\boldsymbol{u} = c(s_1\boldsymbol{v}_1 + s_2\boldsymbol{v}_2) = (cs_1)\boldsymbol{v}_1 + (cs_2)\boldsymbol{v}_2$. ■

若 $\boldsymbol{v}_1$ 不等于零而 $\boldsymbol{v}_2$ 是 $\boldsymbol{v}_1$ 的倍数，则 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 仅生成通过原点的直线.所以通过原点的直线是子空间的另一个例子. 见图 2-25.

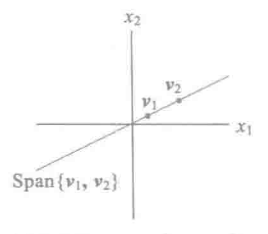


图 2-25 $\boldsymbol{v}_1 \neq 0, \boldsymbol{v}_2 = k\boldsymbol{v}_1$

例 2 不通过原点的一条直线不是子空间，因它不包括原点，同样，图 2-26 说明 L 在加法或标量乘法下不是封闭的.

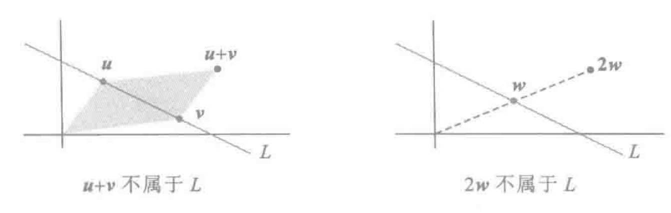


图 2-26 ■

例 3 设 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 属于 $\mathbb{R}^n$，$\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 的所有线性组合是 $\mathbb{R}^n$ 的子空间，这一结论的证明与例 1 中类似，我们将称 $\text{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 为由 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ **生成（或张成）的子空间**. ■

注意 $\mathbb{R}^n$ 是它本身的子空间，因为三个性质都满足. 另一个特殊的子空间是仅含零向量的集合，它也满足子空间的条件，称为**零子空间**.

矩阵的列空间与零空间

应用中，$\mathbb{R}^n$ 的子空间通常出现在以下两种情况中，它们都与矩阵有关.

定义 矩阵 A 的**列空间**是 A 的各列的线性组合的集合，记作 Col A.

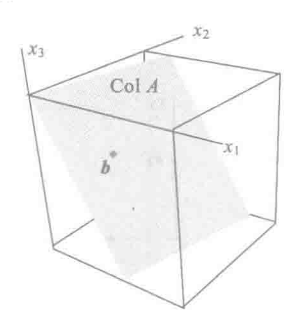
若 $A=[\boldsymbol{a}_1\cdots\boldsymbol{a}_n]$，它们各列属于 $\mathbb{R}^m$，则 Col A 和 $\text{Span}\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n\}$ 相同，例 4 说明 $m\times n$ 矩阵的列空间是 $\mathbb{R}^m$ 的子空间. 注意，仅当 A 的列生成 $\mathbb{R}^m$ 时，Col A 等于 $\mathbb{R}^m$. 否则，Col A 仅是 $\mathbb{R}^m$ 的一部分.

例 4 设 $A=\begin{bmatrix}1&-3&-4\\-4&6&-2\\-3&7&6\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}3\\3\\-4\end{bmatrix}$，确定 $\boldsymbol{b}$ 是否属于 A 的列空间.

解 向量 $\boldsymbol{b}$ 是 A 的各列的线性组合，当且仅当 $\boldsymbol{b}$ 可写成 $A\boldsymbol{x}$ 的形式，$\boldsymbol{x}$ 属于 $\mathbb{R}^3$，也就是说，当且仅当方程 $A\boldsymbol{x}=\boldsymbol{b}$ 有解. 把增广矩阵 $[A\ \ \boldsymbol{b}]$ 进行行变换：

$$\begin{bmatrix}1&-3&-4&3\\-4&6&-2&3\\-3&7&6&-4\end{bmatrix}\sim\begin{bmatrix}1&-3&-4&3\\0&-6&-18&15\\0&-2&-6&5\end{bmatrix}\sim\begin{bmatrix}1&-3&-4&3\\0&-6&-18&15\\0&0&0&0\end{bmatrix}$$

可知 $A\boldsymbol{x}=\boldsymbol{b}$ 相容，从而 $\boldsymbol{b}$ 属于 Col A.



■

例 4 的解答说明，当线性方程组写成 $A\boldsymbol{x}=\boldsymbol{b}$ 的形式，A 的列空间是所有使方程有解的向量 $\boldsymbol{b}$ 的集合.

定义 矩阵 A 的**零空间**是齐次方程 $A\boldsymbol{x}=\boldsymbol{0}$ 的所有解的集合，记为 Nul A.

当 A 有 n 列时，$A\boldsymbol{x}=\boldsymbol{0}$ 的解属于 $\mathbb{R}^n$，A 的零空间是 $\mathbb{R}^n$ 的子集.事实上，Nul A 具有 $\mathbb{R}^n$ 的子空间的性质.

定理 12 $m\times n$ 矩阵 A 的零空间是 $\mathbb{R}^n$ 的子空间. 等价地，n 个未知数的 m 个齐次线性方程的解的全体是 $\mathbb{R}^n$ 的子空间.

证 零向量属于 Nul A（因 $A\boldsymbol{0}=\boldsymbol{0}$），为证明 Nul A 满足其他两个性质，取 Nul A 中两个向量

$\boldsymbol{u}$ 和 $\boldsymbol{v}$，即设 $A\boldsymbol{u}=\boldsymbol{0}$ 和 $A\boldsymbol{v}=\boldsymbol{0}$，那么由矩阵乘法的性质，

$$A(\boldsymbol{u}+\boldsymbol{v})=A\boldsymbol{u}+A\boldsymbol{v}=\boldsymbol{0}+\boldsymbol{0}=\boldsymbol{0}$$

于是 $\boldsymbol{u}+\boldsymbol{v}$ 满足 $A\boldsymbol{x}=\boldsymbol{0}$，所以 $\boldsymbol{u}+\boldsymbol{v}$ 属于 Nul A，同样，对任意数 c，$A(c\boldsymbol{u})=c(A\boldsymbol{u})=c(\boldsymbol{0})=\boldsymbol{0}$，这就证明了 $c\boldsymbol{u}$ 也属于 Nul A. ■

为检验给定向量 $\boldsymbol{v}$ 是否属于 Nul A，只要计算 $A\boldsymbol{v}$，看它是否零向量，因 Nul A 是用其中每个向量必须满足的一个条件来描述的，我们说零空间是隐性定义的. 相反，列空间是显性定义的，因 Col A 中的向量可由 A 的各列（利用线性组合）构造出来. 为了建立 Nul A 的显性描述，解 Ax=0 这个方程，把解写成参数向量形式.（见下面的例 6.）[㊀]

子空间的基

因为子空间一般含有无穷多个向量，子空间中的问题最好能够通过研究生成这个子空间的一个小的有限集合来解决，这个集合越小越好.可以证明，最小可能的生成集合必是线性无关的.

定义　$\mathbb{R}^n$ 中子空间 H 的一组基是 H 中一个线性无关集，它生成 H.

例 5　可逆 $n\times n$ 矩阵的各列构成 $\mathbb{R}^n$ 的一组基. 因为它们线性无关，而且生成 $\mathbb{R}^n$，这由逆矩阵定理可知. 一个这样的矩阵是 $n\times n$ 单位矩阵，它的各列用 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n$ 表示：

$$\boldsymbol{e}_1=\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix},\ \boldsymbol{e}_2=\begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix},\ \cdots,\ \boldsymbol{e}_n=\begin{bmatrix}0\\\vdots\\0\\1\end{bmatrix}$$

$\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$ 称为 $\mathbb{R}^n$ 的**标准基**. 见图 2-27.

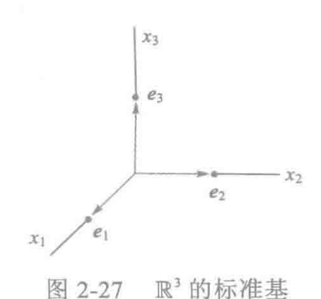


图 2-27　$\mathbb{R}^3$ 的标准基 ■

下例说明，求出方程 $A\boldsymbol{x}=\boldsymbol{0}$ 的解的参数向量形式实际上就是确定 Nul A 的基，这一事实将在第 5 章应用.

例 6　求出下列矩阵的零空间的基.

$$A=\begin{bmatrix}-3&6&-1&1&-7\\1&-2&2&3&-1\\2&-4&5&8&-4\end{bmatrix}$$

㊀ Nul A 与 Col A 的区别将在 4.2 节讨论.

解 首先把方程 $A\boldsymbol{x}=\boldsymbol{0}$ 的解写成参数向量形式.

$$A\sim\begin{bmatrix}1&-2&0&-1&3&0\\0&0&1&2&-2&0\\0&0&0&0&0&0\end{bmatrix},\quad \begin{aligned}x_1-2x_2\quad\ -x_4+3x_5&=0\\x_3+2x_4-2x_5&=0\\0&=0\end{aligned}$$

通解为 $x_1=2x_2+x_4-3x_5,x_3=-2x_4+2x_5,x_2,x_4,x_5$ 为自由变量.

$$\begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{bmatrix}=\begin{bmatrix}2x_2+x_4-3x_5\\x_2\\-2x_4+2x_5\\x_4\\x_5\end{bmatrix}=x_2\underset{\overset{\uparrow}{\boldsymbol{u}}}{\begin{bmatrix}2\\1\\0\\0\\0\end{bmatrix}}+x_4\underset{\overset{\uparrow}{\boldsymbol{v}}}{\begin{bmatrix}1\\0\\-2\\1\\0\end{bmatrix}}+x_5\underset{\overset{\uparrow}{\boldsymbol{w}}}{\begin{bmatrix}-3\\0\\2\\0\\1\end{bmatrix}}\tag{1}$$

$$=x_2\boldsymbol{u}+x_4\boldsymbol{v}+x_5\boldsymbol{w}$$

方程（1）说明 Nul A 与 $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}$ 的所有线性组合的集合是一致的，即 $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$ 生成 Nul A，事实上，$\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}$ 的构造保证了它们线性无关，因为（1）式说明，

$$\boldsymbol{0}=x_2\boldsymbol{u}+x_4\boldsymbol{v}+x_5\boldsymbol{w}$$

仅当权 x_2,x_4,x_5 等于零则成立.（观察向量 $x_2\boldsymbol{u}+x_4\boldsymbol{v}+x_5\boldsymbol{w}$ 的第 2，4，5 个元素.）因此 $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$ 是 Nul A 的一组基. ■

求矩阵的列空间的基比求零空间的基容易. 然而，这个方法需要一些说明，让我们从一个简单情形开始.

例 7 求下列矩阵的列空间的基：

$$B=\begin{bmatrix}1&0&-3&5&0\\0&1&2&-1&0\\0&0&0&0&1\\0&0&0&0&0\end{bmatrix}$$

解 用 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_5$ 表示 B 的列，注意 $\boldsymbol{b}_3=-3\boldsymbol{b}_1+2\boldsymbol{b}_2,\boldsymbol{b}_4=5\boldsymbol{b}_1-\boldsymbol{b}_2,\boldsymbol{b}_3$ 和 $\boldsymbol{b}_4$ 是主元列的线性组合，这意味着 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_5$ 的任意线性组合实际上仅是 $\boldsymbol{b}_1,\boldsymbol{b}_2$ 和 $\boldsymbol{b}_5$ 的线性组合. 事实上，若 $\boldsymbol{v}$ 是 Col B 的任意向量，例如说，

$$\boldsymbol{v}=c_1\boldsymbol{b}_1+c_2\boldsymbol{b}_2+c_3\boldsymbol{b}_3+c_4\boldsymbol{b}_4+c_5\boldsymbol{b}_5$$

则把 $\boldsymbol{b}_3$ 和 $\boldsymbol{b}_4$ 代入，可把 $\boldsymbol{v}$ 写成

$$\boldsymbol{v}=c_1\boldsymbol{b}_1+c_2\boldsymbol{b}_2+c_3(-3\boldsymbol{b}_1+2\boldsymbol{b}_2)+c_4(5\boldsymbol{b}_1-\boldsymbol{b}_2)+c_5\boldsymbol{b}_5$$

它是 $\boldsymbol{b}_1,\boldsymbol{b}_2$ 和 $\boldsymbol{b}_5$ 的线性组合，所以 $\{\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3\}$ 生成 Col B，又 $\boldsymbol{b}_1,\boldsymbol{b}_2$ 和 $\boldsymbol{b}_5$ 为线性无关，因为它们是单位矩阵的列，所以 B 的主元列构成 Col B 的基. ■

例 7 中的矩阵 B 是简化阶梯形.为处理一个一般的矩阵 A，注意 A 的各列之间的线性相关关系可表示为形式 $A\boldsymbol{x}=\boldsymbol{0}$（若某些列不含在特殊的线性相关关系中，则 $\boldsymbol{x}$ 的对应元素为零）. 当 A 行化简为阶梯形 B 时，它的列虽然改变，但方程 $A\boldsymbol{x}=\boldsymbol{0}$ 和 $B\boldsymbol{x}=\boldsymbol{0}$ 有相同的解集. 即 A 的列与 B 的列有相同的线性相关关系.

例 8 可以证明矩阵

$$A=[a_1\ \ a_2\ \ \cdots\ \ a_5]=\begin{bmatrix}1&3&3&2&-9\\-2&-2&2&-8&2\\2&3&0&7&1\\3&4&-1&11&-8\end{bmatrix}$$

行等价于例 7 的矩阵 B，求 Col A 的一组基.

解 由例 7，A 的主元列是第 1，2，5 列. 同时有 $\boldsymbol{b}_3=-3\boldsymbol{b}_1+2\boldsymbol{b}_2$ 和 $\boldsymbol{b}_4=5\boldsymbol{b}_1-\boldsymbol{b}_2$，因行变换不影响列的线性相关关系，有

$$\boldsymbol{a}_3=-3\boldsymbol{a}_1+2\boldsymbol{a}_2 \text{ 和 } \boldsymbol{a}_4=5\boldsymbol{a}_1-\boldsymbol{a}_2$$

经验证这是成立的！由例 7 的讨论，生成 A 的列空间不需要用 $\boldsymbol{a}_3$ 和 $\boldsymbol{a}_4$. 同样，$\{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_5\}$ 必是线性无关的，因 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_5$ 之间的任意线性相关关系必然也使 $\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_5$ 有同样的关系，因 $\{\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_5\}$ 为线性无关，$\{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_5\}$ 也线性无关，因此是 Col A 的一组基. ■

例 8 的讨论可以用来证明下列定理.

定理 13 矩阵 A 的主元列构成列空间的基.

警告 小心，要用 A 的主元列本身作为 Col A 的基，阶梯形 B 的列本身通常并不在 A 的列空间内.（例如在例 7 和例 8 中，B 的列的最后一行都是零，不可能生成 A 的列空间.）

练习题

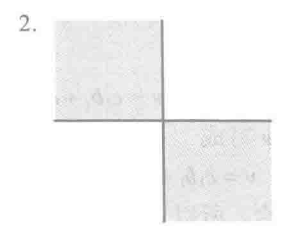
1. 设 $A=\begin{bmatrix}1&-1&5\\2&0&7\\-3&-5&-3\end{bmatrix}$，$\boldsymbol{u}=\begin{bmatrix}-7\\3\\2\end{bmatrix}$，$\boldsymbol{u}$ 是否属于 Nul A？$\boldsymbol{u}$ 是否属于 Col A？给出理由.
2. 设 $A=\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}$，求 Nul A 中的一个向量和 Col A 的一个向量.
3. 设一个 $n\times n$ 矩阵 A 是可逆的，则 Nul A 和 Col A 会如何？

习题 2.8

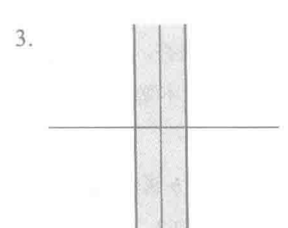
习题 1~4 画的是 $\mathbb{R}^2$ 中的子集，假设这些集合包含边界线，说明这些集 H 不是 $\mathbb{R}^2$ 中的子空间的理由.（例如，找出 H 中两个向量，它们的和不属于 H，或求出 H 中一个向量，它的倍数不属于 H，画出图形.）

1.

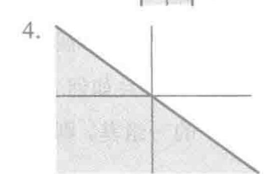
2.



3.



4.



5. $v_1=\begin{bmatrix}1\\3\\-4\end{bmatrix}, v_2=\begin{bmatrix}-2\\-3\\7\end{bmatrix}, w=\begin{bmatrix}-3\\-3\\10\end{bmatrix}$，确定 w 是否属于 $\mathbb{R}^3$ 的由 v_1 和 v_2 生成的子空间？

6. 设 $v_1=\begin{bmatrix}1\\-3\\2\\3\end{bmatrix}, v_2=\begin{bmatrix}4\\-4\\5\\7\end{bmatrix}, v_3=\begin{bmatrix}5\\-3\\6\\5\end{bmatrix}, u=\begin{bmatrix}-1\\-7\\-1\\2\end{bmatrix}$，确定 u 是否属于 $\mathbb{R}^4$ 中由 $\{v_1,v_2,v_3\}$ 生成的子空间？

7. $v_1=\begin{bmatrix}2\\-8\\6\end{bmatrix}, v_2=\begin{bmatrix}-3\\8\\-7\end{bmatrix}, v_3=\begin{bmatrix}-4\\6\\-7\end{bmatrix}, p=\begin{bmatrix}6\\-10\\11\end{bmatrix}, A=[v_1\ v_2\ v_3]$.

a. $\{v_1,v_2,v_3\}$ 中有多少个向量？

b. Col A 中有多少个向量？

c. p 是否属于 Col A？为什么？

8. 设 $v_1=\begin{bmatrix}-2\\0\\6\end{bmatrix}, v_2=\begin{bmatrix}-2\\3\\3\end{bmatrix}, v_3=\begin{bmatrix}0\\-5\\5\end{bmatrix}, p=\begin{bmatrix}-6\\1\\17\end{bmatrix}$，$p$ 是否属于 Col A，其中 $A=[v_1\ \ v_2\ \ v_3]$？

9. A 和 p 如习题 7，判断 p 是否属于 Nul A.

10. $u=\begin{bmatrix}-5\\5\\3\end{bmatrix}$，$A$ 在习题 8 中给出，确定 u 是否属于 Col A.

习题 11~12 给出整数 p 和 q 使 Nul A 是 $\mathbb{R}^p$ 的子空间，Col A 是 $\mathbb{R}^q$ 的子空间.

11. $A=\begin{bmatrix}3&2&1&-5\\-9&-4&1&7\\9&2&-5&1\end{bmatrix}$

12. $A=\begin{bmatrix}1&2&3\\4&5&7\\-5&-1&0\\2&7&11\\3&3&4\end{bmatrix}$

13. A 如习题 11，求 Nul A 和 Col A 中的一个非零向量.

14. A 如习题 12，求 Nul A 和 Col A 中的一个非零向量.

习题 15~20 中，确定哪个集是 $\mathbb{R}^2$ 或 $\mathbb{R}^3$ 中的基？验证你的答案.

15. $\begin{bmatrix}4\\-2\end{bmatrix},\begin{bmatrix}16\\-3\end{bmatrix}$

16. $\begin{bmatrix}-2\\5\end{bmatrix},\begin{bmatrix}4\\-10\end{bmatrix}$

17. $\begin{bmatrix}0\\0\\-2\end{bmatrix},\begin{bmatrix}5\\0\\4\end{bmatrix},\begin{bmatrix}6\\3\\5\end{bmatrix}$

18. $\begin{bmatrix}1\\1\\-3\end{bmatrix},\begin{bmatrix}3\\-1\\2\end{bmatrix},\begin{bmatrix}5\\1\\-4\end{bmatrix}$

19. $\begin{bmatrix}3\\-8\\1\end{bmatrix},\begin{bmatrix}6\\2\\-5\end{bmatrix}$

20. $\begin{bmatrix}1\\-6\\-7\end{bmatrix},\begin{bmatrix}3\\-6\\7\end{bmatrix},\begin{bmatrix}-3\\7\\5\end{bmatrix},\begin{bmatrix}0\\7\\9\end{bmatrix}$

在习题 21 和习题 22 中，标出每个命题的真假.并验证你的答案.

21. a. $\mathbb{R}^n$ 的一个子空间是任一集 H 满足（i）零向量属于 H，（ii）u,v和$u+v$ 属于 H，（iii）c 是数，cu 属于 H.

b. 若 $v_1,\cdots,v_p$ 属于 $\mathbb{R}^n$，则 $\mathrm{Span}\{v_1,\cdots,v_p\}$ 等价于矩阵 $[v_1\cdots v_p]$ 的列空间.

c. 一个 m 行齐次线性方程组有 n 个未知量，其全体解组成的集合是 $\mathbb{R}^m$ 的一个子空间.

d. 一个 $n\times n$ 可逆矩阵的列构成 $\mathbb{R}^n$ 的一组基.

e. 行变换不改变矩阵的列之间的线性相关关系.

22. a. $\mathbb{R}^n$ 中的子集 H 是一个子空间，若零向量属于 H.

b. 若 B 是矩阵 A 的阶梯形，则 B 的主元列构成 Col A 的一组基.

c. 给定 $\mathbb{R}^n$ 中的向量 $v_1,\cdots,v_p$，这些向量的所有线性组合所构成的集合是 $\mathbb{R}^n$ 的一个子空间.

d. 设 H 是 $\mathbb{R}^n$ 的子空间，若 x 属于 H，y 属于 $\mathbb{R}^n$，则 $x+y$ 属于 H.

e. 矩阵 A 的列空间是 $A\boldsymbol{x}=\boldsymbol{b}$ 的解的集合.

习题 23~26 中，给出矩阵 A 和 A 的阶梯形，求出 Col A 和 Nul A 的基.

23. $A=\begin{bmatrix}4&5&9&-2\\6&5&1&12\\3&4&8&-3\end{bmatrix}\sim\begin{bmatrix}1&2&6&-5\\0&1&5&-6\\0&0&0&0\end{bmatrix}$

24. $A=\begin{bmatrix}-3&-6&9&0\\2&-4&7&2\\3&-6&6&-6\end{bmatrix}\sim\begin{bmatrix}1&-2&5&4\\0&0&3&6\\0&0&0&0\end{bmatrix}$

25. $A=\begin{bmatrix}1&4&8&-3&-7\\-1&2&7&3&4\\-2&2&9&5&5\\3&6&9&-5&-2\end{bmatrix}\sim\begin{bmatrix}1&4&8&0&5\\0&2&5&0&-1\\0&0&0&1&4\\0&0&0&0&0\end{bmatrix}$

26. $A=\begin{bmatrix}3&-1&-3&-1&8\\3&1&3&0&2\\0&3&9&-1&-4\\6&3&9&-2&6\end{bmatrix}\sim\begin{bmatrix}3&-1&-3&0&6\\0&2&6&0&-4\\0&0&0&-1&2\\0&0&0&0&0\end{bmatrix}$

27. 构造一个 3×3 矩阵 A 和一个非零向量 $\boldsymbol{b}$，其中 $\boldsymbol{b}$ 属于 Col A 但与 A 的任一列都不同.

28. 构造一个 3×3 矩阵 A 和一个非零向量 $\boldsymbol{b}$，其中 $\boldsymbol{b}$ 不属于 Col A.

29. 构造一个 3×3 矩阵 A 和一个非零向量 $\boldsymbol{b}$，其中 $\boldsymbol{b}$ 属于 Nul A.

30. 设矩阵 $A=[\boldsymbol{a}_1\ \cdots\ \boldsymbol{a}_p]$ 的各列线性无关,说明为什么 $\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_p\}$ 是 Col A 的一组基.

习题 31~36 中,尽可能全面地回答，给出理由.

31. 设 F 是 5×5 矩阵，其列空间不等于 $\mathbb{R}^5$，则 Nul F 会如何？

32. 若 B 是 7×7 矩阵，Col $B=\mathbb{R}^7$，$\boldsymbol{b}$ 属于 $\mathbb{R}^7$,则对形如 $B\boldsymbol{x}=\boldsymbol{b}$ 的方程，其解会如何？

33. 若 C 是 6×6 矩阵，Nul C 是零子空间,$\boldsymbol{b}$ 属于 $\mathbb{R}^6$，则对形如 $C\boldsymbol{x}=\boldsymbol{b}$ 的方程，其解会如何？

34. 若 $m\times n$ 矩阵 A 的各列构成 $\mathbb{R}^m$ 的一组基，则该矩阵的形状会如何？

35. 若 B 是 5×5 矩阵，Nul B 不是零子空间，则 Col B 会如何？

36. 若 C 是 6×4 矩阵，有线性无关的列，则 Nul C 会如何？

[M]习题 37 和习题 38 中，求出给定矩阵 A 的列空间和零空间.

37. $A=\begin{bmatrix}3&-5&0&-1&3\\-7&9&-4&9&-11\\-5&7&-2&5&-7\\3&-7&-3&4&0\end{bmatrix}$

38. $A=\begin{bmatrix}5&3&2&-6&-8\\4&1&3&-8&-7\\5&1&4&5&19\\-7&-5&-2&8&5\end{bmatrix}$

练习题答案

1. 为确定 $\boldsymbol{u}$ 是否属于 Nul A，只需计算

$$A\boldsymbol{u}=\begin{bmatrix}1&-1&5\\2&0&7\\-3&-5&-3\end{bmatrix}\begin{bmatrix}-7\\3\\2\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

结果显示 $\boldsymbol{u}$ 是属于 Nul A. 要确定 $\boldsymbol{u}$ 是否属于 Col A 需要做更多的工作. 化简增广矩阵 $[A\ \ \boldsymbol{u}]$ 成阶梯形来判断方程 $A\boldsymbol{x}=\boldsymbol{u}$ 是否相容：

$$\begin{bmatrix}1&-1&5&-7\\2&0&7&3\\-3&-5&-3&2\end{bmatrix}\sim\begin{bmatrix}1&-1&5&-7\\0&2&-3&17\\0&-8&12&-19\end{bmatrix}\sim\begin{bmatrix}1&-1&5&-7\\0&2&-3&17\\0&0&0&49\end{bmatrix}$$

方程 $A\boldsymbol{x}=\boldsymbol{u}$ 无解，因此 $\boldsymbol{u}$ 不属于 Col A.

2. 与练习题 1 相比，求 Nul A 的一个向量比判定一个给定的向量是否属于 Nul A 需要做更多的工作. 但是，由于 A 已经是阶梯形，方程 $A\boldsymbol{x}=\boldsymbol{0}$ 给出如果 $\boldsymbol{x}=(x_1,x_2,x_3)$，则 $x_2=0, x_3=0, x_1$ 是自由变量. 因此，Nul A 的一个基是 $\boldsymbol{v}=(1,0,0)$. 求 Col A 的一个向量是简单的，因为 A 中的每一列都属于 Col A. 此外，同一个向量 $\boldsymbol{v}$ 既属于 Nul A，又属于 Col A. 对大部分矩阵而言，只有零向量是既属于 Nul A，又属于 Col A.

3. 如果 A 是可逆的，则根据可逆矩阵定理，A 的各列生成 $\mathbb{R}^n$. 由定义可知，任何矩阵的列总是可以生成该矩阵的列空间，因此这里 Col A 就是全部的 $\mathbb{R}^n$. 用符号表示就是 Col $A=\mathbb{R}^n$. 同时，因为 A 是可逆的，方程 $A\boldsymbol{x}=\boldsymbol{0}$ 只有平凡解. 这意味着 Nul A 是零子空间,用符号表示就是 Nul $A=\{\boldsymbol{0}\}$.

2.9 维数与秩

本节从坐标系的概念开始对子空间和子空间的基继续加以讨论. 下面的定义和例子使一个有用的新术语——维数，显得非常自然，至少对 $\mathbb{R}^3$ 子空间是这样.

坐标系

选择子空间 H 的一个基代替一个纯粹生成集的主要原因，是 H 中的每个向量可以被表示为基向量的线性组合的唯一形式. 为了明确原因，假设 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_p\}$ 是 H 的基，H 中的一个向量 $\boldsymbol{x}$ 可以由两种方式生成，设

$$\boldsymbol{x}=c_1\boldsymbol{b}_1+\cdots+c_p\boldsymbol{b}_p \text{，且 } \boldsymbol{x}=d_1\boldsymbol{b}_1+\cdots+d_p\boldsymbol{b}_p \tag{1}$$

则相减得到

$$\boldsymbol{0}=\boldsymbol{x}-\boldsymbol{x}=(c_1-d_1)\boldsymbol{b}_1+\cdots+(c_p-d_p)\boldsymbol{b}_p \tag{2}$$

因为 $\mathcal{B}$ 是线性无关的，（2）中的权值必全为零. 亦即对 $1\leqslant j\leqslant p$, $c_j=d_j$，（1）式中的两种表示实际上是相同的.

定义 假设 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_p\}$ 是子空间 H 的一组基，对 H 中的每一个向量 $\boldsymbol{x}$，相对于基 $\mathcal{B}$ 的坐标是使 $\boldsymbol{x}=c_1\boldsymbol{b}_1+\cdots+c_p\boldsymbol{b}_p$ 成立的权值 $c_1,\cdots,c_p$，且 $\mathbb{R}^p$ 中的向量

$$[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}$$

称为 $\boldsymbol{x}$（相对于 $\mathcal{B}$）的坐标向量，或 $\boldsymbol{x}$ 的 $\mathcal{B}$-坐标向量.[㊀]

㊀ 要注意 $\mathcal{B}$ 中向量的次序，因为 $[\boldsymbol{x}]_{\mathcal{B}}$ 的元素依赖于 $\mathcal{B}$ 中向量的次序.

例 1 设 $\boldsymbol{v}_1=\begin{bmatrix}3\\6\\2\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}-1\\0\\1\end{bmatrix},\boldsymbol{x}=\begin{bmatrix}3\\12\\7\end{bmatrix},\mathcal{B}=\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$. 则因 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 线性无关，$\mathcal{B}$ 是 $H=\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 的基. 判断 $\boldsymbol{x}$ 是否在 H 中，如果是，求 $\boldsymbol{x}$ 相对基 $\mathcal{B}$ 的坐标向量.

解 如果 $\boldsymbol{x}$ 在 H 中，则下面的向量方程是相容的：

$$c_1\begin{bmatrix}3\\6\\2\end{bmatrix}+c_2\begin{bmatrix}-1\\0\\1\end{bmatrix}=\begin{bmatrix}3\\12\\7\end{bmatrix}$$

如果数 c_1,c_2 存在，即是 $\boldsymbol{x}$ 的 $\mathcal{B}$-坐标. 由行操作得

$$\begin{bmatrix}3&-1&3\\6&0&12\\2&1&7\end{bmatrix}\sim\begin{bmatrix}1&0&2\\0&1&3\\0&0&0\end{bmatrix}$$

于是 $c_1=2,c_2=3,[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}2\\3\end{bmatrix}$. 基 $\mathcal{B}$ 确定 H 上的一个"坐标系",如图 2-28 中的格子所示. ■

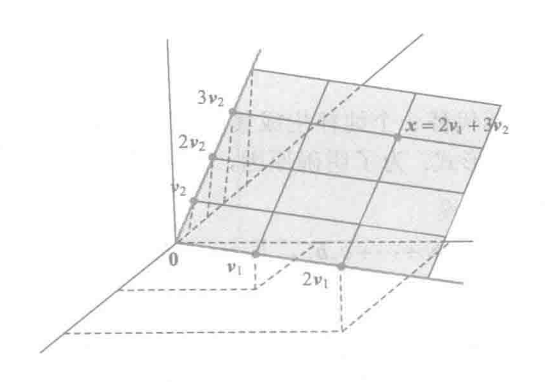


图 2-28 $\mathbb{R}^3$ 中一个平面 H 的一个坐标系

注意到虽然 H 中的点也在 $\mathbb{R}^3$ 中，它们完全由属于 $\mathbb{R}^2$ 中的坐标向量确定. 图 2-28 中的平面上的格子使 H 看起来像 $\mathbb{R}^2$. 映射 $\boldsymbol{x}\mapsto[\boldsymbol{x}]_{\mathcal{B}}$ 是 H 和 $\mathbb{R}^2$ 之间保持线性组合关系的一一映射. 我们称这种映射是同构的，且 H 与 $\mathbb{R}^2$ 同构.

一般地，如果 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_p\}$ 是 H 的基，则映射 $\boldsymbol{x}\mapsto[\boldsymbol{x}]_{\mathcal{B}}$ 是使 H 和 $\mathbb{R}^p$ 的形态一样的一一映射（尽管 H 中的向量可能有多于 p 个元素）.（详细的讨论见 4.4 节.）

子空间的维数

可以证明，若子空间 H 有一组基包含 p 个向量，则 H 的每个基都正好包含 p 个向量，（见习题 27 和 28），于是下列定义是有意义的.

定义 非零子空间 H 的**维数**，用 $\dim H$ 表示，是 H 的任意一个基的向量个数. 零子空间

$\{\mathbf{0}\}$ 的维数定义为零.[1]

$\mathbb{R}^n$ 空间维数为 n，$\mathbb{R}^n$ 的每个基由 n 个向量组成. $\mathbb{R}^3$ 中一个经过 $\mathbf{0}$ 的平面是 2 维的，一条经过 $\mathbf{0}$ 的直线是一维的.

例 2 回忆 2.8 节例 6 中矩阵 A 的零空间有一个基包含 3 个向量. 因此这里 Nul A 的维数为 3. 观察到每个基向量对应方程 $A\mathbf{x}=\mathbf{0}$ 的一个自由变量. 我们的构造方法总是以这种方式产生一个基. 因此，要确定 Nul A 的维数，只需求出 $A\mathbf{x}=\mathbf{0}$ 中的自由变量个数. ■

定义 矩阵 A 的秩（记为 rank A）是 A 的列空间的维数.

因为 A 的主元列形成 Col A 的一个基，A 的秩正好是 A 的主元列的个数.

例 3 确定矩阵的秩

$$A=\begin{bmatrix}2 & 5 & -3 & -4 & 8\\ 4 & 7 & -4 & -3 & 9\\ 6 & 9 & -5 & 2 & 4\\ 0 & -9 & 6 & 5 & -6\end{bmatrix}$$

解 简化 A 成阶梯行：

$$A\sim\begin{bmatrix}2 & 5 & -3 & -4 & 8\\ 0 & -3 & 2 & 5 & -7\\ 0 & -6 & 4 & 14 & -20\\ 0 & -9 & 6 & 5 & -6\end{bmatrix}\sim\cdots\sim\begin{bmatrix}2 & 5 & -3 & -4 & 8\\ 0 & -3 & 2 & 5 & -7\\ 0 & 0 & 0 & 4 & -6\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}$$

主元列（第 1、2、4 列）

矩阵 A 有 3 个主元列，因此 rank $A=3$. ■

从例 3 的行化简可见在 $A\mathbf{x}=\mathbf{0}$ 中有两个自由变量，因为 A 的五列中有两列不是主元列.（非主元列对应于 $A\mathbf{x}=\mathbf{0}$ 中的自由变量.）由于主元列的个数加上非主元列的个数正好是 A 的列数，Col A 和 Nul A 的维数有如下有用的关系.（详细内容见 4.6 节的秩定理.）

定理 14 （秩定理）

如果一矩阵 A 有 n 列，则 rank A + dim Nul $A=n$.

下面的定理在应用中很重要，并在第 5 章和第 6 章中用到. 该定理（在 4.5 节中证明）当然是很显明的，若你想到 p 维子空间同构于 $\mathbb{R}^p$. 由可逆矩阵定理知，$\mathbb{R}^p$ 中的 p 个向量线性无关，当且仅当这 p 个向量也生成 $\mathbb{R}^p$.

定理 15 （基定理）

设 H 是 $\mathbb{R}^n$ 的 p 维子空间，H 中的任何恰好由 p 个成员组成的线性无关集构成 H 的一个基. 并且，H 中任何生成 H 的 p 个向量集也构成 H 的一个基.

---

[1] 零子空间无基（因为零向量本身构成一个线性相关集）.

秩与可逆矩阵定理

各种与矩阵相关的向量空间的概念为可逆矩阵定理提供了更多的命题. 下面给出 2.3 节原定理的后续命题.

定理 （可逆矩阵定理（续））

设 A 是一 $n\times n$ 矩阵，则下面的每个命题与 A 是可逆矩阵的命题等价：

m. A 的列向量构成 $\mathbb{R}^n$ 的一个基.

n. $\text{Col}\,A=\mathbb{R}^n$.

o. $\dim\text{Col}\,A=n$.

p. $\text{rank}\,A=n$.

q. $\text{Nul}\,A=\{\mathbf{0}\}$.

r. $\dim\text{Nul}\,A=0$.

证 根据线性无关和生成的概念，命题（m）逻辑上与命题（e）和（h）等价. 其他五个命题通过简单推导以如下关系与定理以前的命题相连：

$$(g)\Rightarrow(n)\Rightarrow(o)\Rightarrow(p)\Rightarrow(r)\Rightarrow(q)\Rightarrow(d)$$

命题（g）认为方程 $A\boldsymbol{x}=\boldsymbol{b}$ 对每一属于 $\mathbb{R}^n$ 的 $\boldsymbol{b}$ 有至少一个解，由此可以推出（n），因为 $\text{Col}\,A$ 确实是所有 $\boldsymbol{b}$ 的集合，满足方程 $A\boldsymbol{x}=\boldsymbol{b}$ 相容的条件. 命题（n）$\Rightarrow$（o）$\Rightarrow$（p）是因为维数和秩的定义. 如果 A 的秩是 n，即 A 的列数，则根据秩定理得 $\dim\text{Nul}\,A=0$，因而 $\text{Nul}\,A=\{\mathbf{0}\}$. 于是有（p）$\Rightarrow$（r）$\Rightarrow$（q）. 同时，由命题（q）推出方程 $A\boldsymbol{x}=\mathbf{0}$ 只有平凡解，即命题（d）. 因为已知命题（d）和（g）与 A 是可逆矩阵的命题等价，从而定理证毕. ■

数值计算的注解 本教材中讨论的许多算法有助于概念的理解和手工进行简单的计算. 然而这些算法通常不适于处理现实生活中的大规模问题.

计算秩的算法是一个很好的例子. 表面上看将矩阵简化为阶梯阵后数主元是很容易的事情. 但除非是对元素精确指定的矩阵进行算术运算，行运算可以明显改变一个矩阵的秩. 例如，假如矩阵 $\begin{bmatrix}5 & 7\\ 5 & x\end{bmatrix}$ 中 x 的值在计算机中不是存为 7，那么它的秩可能是 1 或 2，取决于计算机是否视 $x-7$ 为零.

在实际应用中，通常使用 A 的奇异值分解有效地确定矩阵 A 的秩，在 7.4 节中将会讨论.

练习题

1. 确定由向量 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 生成的 $\mathbb{R}^3$ 的子空间 H 的维数.（首先找 H 的基.）

$$\boldsymbol{v}_1=\begin{bmatrix}2\\-8\\6\end{bmatrix},\ \boldsymbol{v}_2=\begin{bmatrix}3\\-7\\-1\end{bmatrix},\ \boldsymbol{v}_3=\begin{bmatrix}-1\\6\\-7\end{bmatrix}$$

2. $\mathbb{R}^2$ 的基 $\mathcal{B}=\left\{\begin{bmatrix}1\\0.2\end{bmatrix},\begin{bmatrix}0.2\\1\end{bmatrix}\right\}$，若 $[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}3\\2\end{bmatrix}$，$\boldsymbol{x}$ 是什么？

3. $\mathbb{R}^3$ 是否可能包含 4 维子空间？为什么？

习题 2.9

习题 1~2 给出坐标向量 $[\boldsymbol{x}]_{\mathcal{B}}$ 和基 $\mathcal{B}$，求向量 $\boldsymbol{x}$．像练习题 2 一样用图示说明你的答案.

1. $\mathcal{B}=\left\{\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}2\\-1\end{bmatrix}\right\}$，$[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}3\\2\end{bmatrix}$

2. $\mathcal{B}=\left\{\begin{bmatrix}-3\\1\end{bmatrix},\begin{bmatrix}3\\2\end{bmatrix}\right\}$，$[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}-1\\2\end{bmatrix}$

习题 3~6 中，$\boldsymbol{x}$ 属于一个基为 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$ 的子空间 H，求出 $\boldsymbol{x}$ 相对于 $\mathcal{B}$ 的坐标向量.

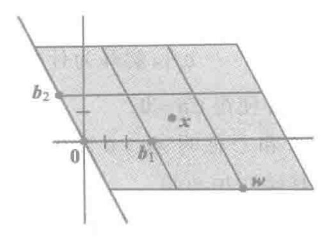
3. $\boldsymbol{b}_1=\begin{bmatrix}2\\-3\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}-1\\5\end{bmatrix},\boldsymbol{x}=\begin{bmatrix}0\\7\end{bmatrix}$

4. $\boldsymbol{b}_1=\begin{bmatrix}1\\-5\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}-2\\3\end{bmatrix},\boldsymbol{x}=\begin{bmatrix}-1\\9\end{bmatrix}$

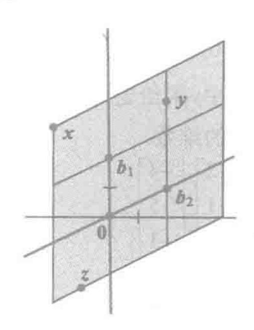
5. $\boldsymbol{b}_1=\begin{bmatrix}1\\4\\-3\end{bmatrix}$, $\boldsymbol{b}_2=\begin{bmatrix}-2\\-7\\5\end{bmatrix}$, $\boldsymbol{x}=\begin{bmatrix}2\\9\\-7\end{bmatrix}$

6. $\boldsymbol{b}_1=\begin{bmatrix}-3\\2\\-4\end{bmatrix}$, $\boldsymbol{b}_2=\begin{bmatrix}7\\-3\\5\end{bmatrix}$, $\boldsymbol{x}=\begin{bmatrix}5\\0\\-2\end{bmatrix}$

7. 设 $\boldsymbol{b}_1=\begin{bmatrix}3\\0\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}-1\\2\end{bmatrix},\boldsymbol{w}=\begin{bmatrix}7\\-2\end{bmatrix},\boldsymbol{x}=\begin{bmatrix}4\\1\end{bmatrix},\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$．使用图形估算 $[\boldsymbol{w}]_{\mathcal{B}}$ 和 $[\boldsymbol{x}]_{\mathcal{B}}$．用你估算的 $[\boldsymbol{x}]_{\mathcal{B}}$ 和 $\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$ 计算 $\boldsymbol{x}$ 来验证 $[\boldsymbol{x}]_{\mathcal{B}}$.



8. 设 $\boldsymbol{b}_1=\begin{bmatrix}0\\2\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}2\\1\end{bmatrix},\boldsymbol{x}=\begin{bmatrix}-2\\3\end{bmatrix},\boldsymbol{y}=\begin{bmatrix}2\\4\end{bmatrix},\boldsymbol{z}=\begin{bmatrix}-1\\-2.5\end{bmatrix}$，$\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$．使用图形估算 $[\boldsymbol{x}]_{\mathcal{B}},[\boldsymbol{y}]_{\mathcal{B}}$ 和 $[\boldsymbol{z}]_{\mathcal{B}}$．用你估算的 $[\boldsymbol{y}]_{\mathcal{B}},[\boldsymbol{z}]_{\mathcal{B}}$ 和 $\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$ 计算 $\boldsymbol{y}$ 和 $\boldsymbol{z}$ 来验证 $[\boldsymbol{y}]_{\mathcal{B}}$ 和 $[\boldsymbol{z}]_{\mathcal{B}}$.



习题 9~12 给出矩阵 A 及其阶梯阵，求 Col A 和 Nul A 的基，并求其维数.

9. $A=\begin{bmatrix}1&3&2&-6\\3&9&1&5\\2&6&-1&9\\5&15&0&14\end{bmatrix}\sim\begin{bmatrix}1&3&3&2\\0&0&5&-7\\0&0&0&5\\0&0&0&0\end{bmatrix}$

10. $A=\begin{bmatrix}1&-2&-1&5&4\\2&-1&1&5&6\\-2&0&-2&1&-6\\3&1&4&1&5\end{bmatrix}\sim\begin{bmatrix}1&-2&-1&2&0\\0&1&1&0&3\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix}$

11. $A=\begin{bmatrix}2&4&-5&2&-3\\3&6&-8&3&-5\\0&0&9&0&9\\-3&-6&-7&-3&10\end{bmatrix}\sim\begin{bmatrix}1&2&-5&1&-4\\0&0&5&0&5\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix}$

12. $A=\begin{bmatrix}1&2&-4&4&6\\5&1&-9&2&10\\4&6&-9&12&15\\3&4&-5&8&9\end{bmatrix}$

$\sim\begin{bmatrix}1&2&8&4&6\\0&2&3&4&-1\\0&0&5&0&-5\\0&0&0&0&0\end{bmatrix}$

习题 13~14 中,求给定向量生成的子空间的一个基及该子空间的维数.

13. $\begin{bmatrix}1\\-3\\2\\-4\end{bmatrix},\begin{bmatrix}-3\\9\\-6\\12\end{bmatrix},\begin{bmatrix}2\\-1\\4\\2\end{bmatrix},\begin{bmatrix}-4\\5\\-3\\7\end{bmatrix}$

14. $\begin{bmatrix}1\\-1\\-2\\3\end{bmatrix},\begin{bmatrix}2\\-3\\-1\\4\end{bmatrix},\begin{bmatrix}0\\-1\\3\\2\end{bmatrix},\begin{bmatrix}-1\\4\\-7\\7\end{bmatrix},\begin{bmatrix}3\\-7\\6\\-9\end{bmatrix}$

15. 设一 4×6 矩阵 A 有 4 个主元列. 是否 $\mathrm{Col}\,A=\mathbb{R}^4$? 是否 $\mathrm{Nul}\,A=\mathbb{R}^2$? 解释原因.

16. 设一 4×7 矩阵 A 有三个主元列. 是否 $\mathrm{Col}\,A=\mathbb{R}^3$? $\mathrm{Nul}\,A$ 的维数是多少? 解释原因.

习题 17~18 中，标出每个命题的真假，给出理由. 这里 A 是 $m\times n$ 矩阵.

17. a. 若 $\mathcal{B}=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 是子空间 H 的一个基，$\boldsymbol{x}=c_1\boldsymbol{v}_1+\cdots+c_p\boldsymbol{v}_p$，则 $c_1,\cdots,c_p$ 是 $\boldsymbol{x}$ 相对基 $\mathcal{B}$ 的坐标.

b. $\mathbb{R}^n$ 中的每一直线是 $\mathbb{R}^n$ 的一维子空间.

c. $\mathrm{Col}\,A$ 的维数是 A 的主元列的个数.

d. $\mathrm{Col}\,A$ 和 $\mathrm{Nul}\,A$ 的维数之和等于 A 的列数.

e. 若 p 个向量生成 $\mathbb{R}^n$ 中的 p 维子空间 H，则这 p 个向量构成 H 的基.

18. a. 若 $\mathcal{B}=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 是子空间 H 的一个基，则 H 中的每个向量可以写成 $\mathcal{B}$ 中向量的唯一线性组合形式.

b. $\mathrm{Nul}\,A$ 的维数是方程 $A\boldsymbol{x}=\boldsymbol{0}$ 中的变量个数.

c. A 的列空间的维数是 $\mathrm{rank}\,A$.

d. 若 $\mathcal{B}=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 是 $\mathbb{R}^n$ 中的子空间 H 的基，则映射 $\boldsymbol{x}\mapsto[\boldsymbol{x}]_{\mathcal{B}}$ 使 H 和 $\mathbb{R}^p$ 的形态一样.

e. 若 H 是 $\mathbb{R}^n$ 中的 p 维子空间，则 H 中的 p 个线性无关向量组成的集合构成 H 的基.

在习题 19~24 中，给出回答的证明或构造.

19. 若 $A\boldsymbol{x}=\boldsymbol{0}$ 的解空间有一个由 3 个向量组成的基，A 为 5×7 矩阵，A 的秩是多少?

20. 若 6×8 矩阵的零空间是 3 维的，它的秩是多少?

21. 一个 9×8 矩阵 A 的秩是 7，$A\boldsymbol{x}=\boldsymbol{0}$ 的解空间维数是多少?

22. 证明：若 $\dim\mathrm{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_5\}=4$，则 $\mathbb{R}^n$ 中的 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_5\}$ 线性无关.

23. 构造一个 $\dim\mathrm{Nul}\,A=3$ 和 $\dim\mathrm{Col}\,A=2$ 的 3×5 矩阵.

24. 构造一个秩等于 1 的 3×4 矩阵.

25. 设一 $n\times p$ 矩阵 A 的列空间是 p 维的，解释为什么 A 的列向量一定线性无关.

26. 假设 A 的第 1，3，4，5 和 7 列向量线性无关(但不必是主元列)且 A 的秩是 5，解释为什么这 5 个列向量一定是 A 的列空间的一个基.

27. 假设向量 $\boldsymbol{b}_1,\ldots,\boldsymbol{b}_p$ 生成一个子空间 W，$\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_q\}$ 是任一 W 中包含多于 p 个向量的向量集，完成下面的论述，证明 $\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_q\}$ 一定是线性相关的. 首先记 $B=[\boldsymbol{b}_1\ \cdots\ \boldsymbol{b}_p]$, $A=[\boldsymbol{a}_1\ \cdots\ \boldsymbol{a}_q]$.

a. 解释为什么对每一向量 $\boldsymbol{a}_j$，存在一个 $\mathbb{R}^p$ 中的向量 $\boldsymbol{c}_j$ 使得 $\boldsymbol{a}_j=B\boldsymbol{c}_j$.

b. 设 $C=[\boldsymbol{c}_1\ \cdots\ \boldsymbol{c}_q]$，解释为什么存在一个非零向量 $\boldsymbol{u}$ 使得 $C\boldsymbol{u}=\boldsymbol{0}$.

c. 使用 B 和 C 证明 $A\boldsymbol{u}=\boldsymbol{0}$，从而证明 A 的列向量是线性相关的.

28. 使用习题 27 证明：如果$\mathcal{A}$和$\mathcal{B}$是 $\mathbb{R}^n$ 中一子空间 W 的基，则$\mathcal{A}$不可能包含比$\mathcal{B}$更多的向量，相反地，$\mathcal{B}$也不可能包含比$\mathcal{A}$更多的向量.

29. [M]设 $H=\mathrm{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$, $\mathcal{B}=\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$，证明 $\boldsymbol{x}$ 属于 H，并求 $\boldsymbol{x}$ 相对$\mathcal{B}$的坐标向量.

$$v_1=\begin{bmatrix}15\\-5\\12\\7\end{bmatrix}, v_2=\begin{bmatrix}14\\-10\\13\\17\end{bmatrix}, x=\begin{bmatrix}16\\0\\11\\-3\end{bmatrix}$$

30. [M] 设 $H=\mathrm{Span}\{v_1,v_2,v_3\}$, $\mathcal{B}=\{v_1,v_2,v_3\}$，证明 $\mathcal{B}$ 是 H 的基，x 属于 H，并求 x 相对 $\mathcal{B}$ 的坐标向量.

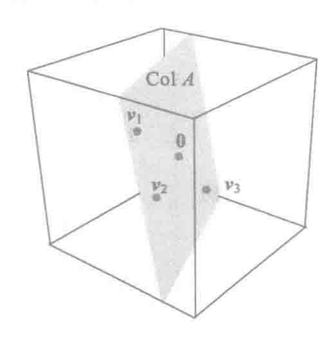
$$v_1=\begin{bmatrix}-6\\3\\-9\\4\end{bmatrix}, v_2=\begin{bmatrix}8\\0\\7\\-3\end{bmatrix}, v_3=\begin{bmatrix}-9\\5\\-8\\3\end{bmatrix}, x=\begin{bmatrix}11\\-2\\17\\-8\end{bmatrix}$$

练习题答案

1. 构造 $A=[v_1\ \ v_2\ \ v_3]$，使由 v_1,v_2,v_3 生成的子空间是 A 的列空间. 这一空间的一组基是 A 的主元列.

$$A=\begin{bmatrix}2&3&-1\\-8&-7&6\\6&-1&-7\end{bmatrix}\sim\begin{bmatrix}2&3&-1\\0&5&2\\0&-10&-4\end{bmatrix}\sim\begin{bmatrix}2&3&-1\\0&5&2\\0&0&0\end{bmatrix}$$

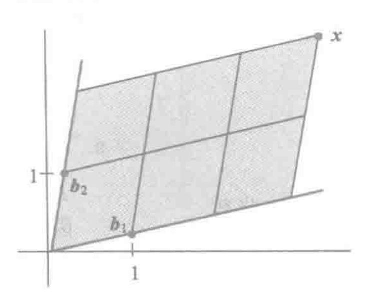
由此 A 的前两列是主元列，构成 H 的基，所以 $\dim H=2$.



2. 若 $[x]_{\mathcal{B}}=\begin{bmatrix}3\\2\end{bmatrix}$，则 x 是基向量的线性组合，权为 3 和 2.

$$x=3b_1+2b_2=3\begin{bmatrix}1\\0.2\end{bmatrix}+2\begin{bmatrix}0.2\\1\end{bmatrix}=\begin{bmatrix}3.4\\2.6\end{bmatrix}$$

基 $\{b_1, b_2\}$ 确定 $\mathbb{R}^2$ 的坐标系，在图中用格子表示. 注意 x 在 b_1 方向是 3 个单位，b_2 方向是 2 个单位.



3. 4 维空间包含一个有 4 个线性无关向量的基，这在 $\mathbb{R}^3$ 中是不可能的. 因 $\mathbb{R}^3$ 中任意线性无关集不超过 3 个向量. $\mathbb{R}^3$ 中任意子空间的维数不超过 3，空间 $\mathbb{R}^3$ 本身是 $\mathbb{R}^3$ 仅有的 3 维子空间. 其他子空间的维数为 2，1 或 0.

补充习题

1. 设在以下命题中提到的矩阵有适当的维数. 标出各个命题的真假，给出理由.
   a. 若 A 和 B 都是 $m\times n$ 矩阵，则 AB^{T} 和 $A^{\mathrm{T}}B$ 都有定义.
   b. 若 $AB=C$ 而 C 有 2 列，则 A 有 2 列.
   c. 将矩阵 B 以对角元素非零的对角矩阵 A 左乘，等于对 B 的各行进行标量乘法.
   d. 若 $BC=BD$，则 $C=D$.
   e. 若 $AC=0$，则 $A=0$ 或 $C=0$.
   f. 若 A、B 都是 $n\times n$ 矩阵，则 $(A+B)(A-B)=A^2-B^2$.
   g. 一个初等 $n\times n$ 矩阵有 n 或 $n+1$ 个非零元素.
   h. 初等矩阵的转置也是初等矩阵.
   i. 初等矩阵必定是方阵.
   j. 每个方阵是初等矩阵的积.
   k. 若 A 是 3×3 矩阵，有 3 个主元位置，存在初等矩阵 $E_1,\cdots,E_p$ 使 $E_p\cdots E_1A=I$.
   l. 若 $AB=I$，则 A 可逆.
   m. 若 A 和 B 是可逆矩阵，则 AB 可逆且 $(AB)^{-1}=A^{-1}B^{-1}$.
   n. 若 $AB=BA$ 且 A 可逆，则 $A^{-1}B=BA^{-1}$.
   o. 若 A 可逆且 $r\neq0$，则 $(rA)^{-1}=rA^{-1}$.
   p. 若 A 是 3×3 矩阵且方程 $A\boldsymbol{x}=\begin{bmatrix}1\\0\\0\end{bmatrix}$ 有唯一解，则 A 可逆.
2. 求矩阵 C 使 $C^{-1}=\begin{bmatrix}4&5\\6&7\end{bmatrix}$.
3. 设 $A=\begin{bmatrix}0&0&0\\1&0&0\\0&1&0\end{bmatrix}$，证明 $A^3=0$，利用矩阵代数计算乘积 $(I-A)(I+A+A^2)$.
4. 设对某个 $n>1$，$A^n=0$，求 $I-A$ 的逆.
5. 设 $n\times n$ 矩阵 A 满足方程 $A^2-2A+I=0$，证明：$A^3=3A-2I$，$A^4=4A-3I$.
6. 设 $A=\begin{bmatrix}1&0\\0&-1\end{bmatrix},B=\begin{bmatrix}0&1\\1&0\end{bmatrix}$，它们是 Pauli spin 矩阵，在量子力学的电子旋转研究中有应用. 证明: $A^2=I$，$B^2=I$，且 $AB=-BA$，矩阵满足 $AB=-BA$ 称为负交换的.
7. 设 $A=\begin{bmatrix}1&3&8\\2&4&11\\1&2&5\end{bmatrix},B=\begin{bmatrix}-3&5\\1&5\\3&4\end{bmatrix}$，不通过计算 A^{-1} 求 $A^{-1}B$.（提示：$A^{-1}B$ 是方程 $AX=B$ 的解.）
8. 求矩阵 A 使变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 把 $\begin{bmatrix}1\\3\end{bmatrix}$ 和 $\begin{bmatrix}2\\7\end{bmatrix}$ 分别变为 $\begin{bmatrix}1\\1\end{bmatrix}$ 和 $\begin{bmatrix}3\\1\end{bmatrix}$.（提示：写个包含矩阵 A 的方程，并把它解出来.）
9. 设 $AB=\begin{bmatrix}5&4\\-2&3\end{bmatrix},B=\begin{bmatrix}7&3\\2&1\end{bmatrix}$，求 A.
10. 设 A 可逆. 说明为什么 $A^{\mathrm{T}}A$ 也可逆，然后证明 $A^{-1}=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$.
11. 设 $x_1,\cdots,x_n$ 为固定的数，下列矩阵称为范德蒙德矩阵，在信号处理、纠错码以及多项式插值中都有应用.
$$V=\begin{bmatrix}1&x_1&x_1^2&\cdots&x_1^{n-1}\\1&x_2&x_2^2&\cdots&x_2^{n-1}\\\vdots&\vdots&\vdots&&\vdots\\1&x_n&x_n^2&\cdots&x_n^{n-1}\end{bmatrix}$$
给定 $\mathbb{R}^n$ 中 $\boldsymbol{y}=(y_1,y_2,\cdots,y_n)$，设 $\mathbb{R}^n$ 中 $\boldsymbol{c}=(c_0,c_1,\cdots,c_{n-1})$ 满足 $V\boldsymbol{c}=\boldsymbol{y}$，定义多项式
$$p(t)=c_0+c_1t+c_2t^2+\cdots+c_{n-1}t^{n-1}$$
   a. 证明 $p(x_1)=y_1,p(x_2)=y_2,\cdots,p(x_n)=y_n$，称 $p(t)$ 为点 $(x_1,y_1),\cdots,(x_n,y_n)$ 的插值多项式，因 $p(t)$ 的图像通过这些点.
   b. 设 $x_1,x_2,\cdots,x_n$ 为相异的数，证明 V 的列是线性无关的.（提示：一个 $n-1$ 次多项式至多能有多少个零点？）
   c. 证明：若 $x_1,\cdots,x_n$ 为相异的数，而 $y_1,\cdots,y_n$ 是

任意数，则存在一个过点 $(x_1,y_1),\cdots,(x_n,y_n)$ 且次数 $\leqslant n-1$ 次的插值多项式.

12. 设 $A=LU$，其中 L 是可逆下三角矩阵，U 是上三角矩阵，说明为什么 A 的第一列是 L 的第一列的倍数. A 的第二列和 L 的列有何关系？

13. 给定 $\boldsymbol{u}$ 属于 $\mathbb{R}^n$，$\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}=1$，设 $P=\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}$（外积），而 $Q=I-2P$，证明：

    a. $P^2=P$　　b. $P^{\mathrm{T}}=P$　　c. $Q^2=I$

    变换 $\boldsymbol{x}\mapsto P\boldsymbol{x}$ 称为一个投影，$\boldsymbol{x}\mapsto Q\boldsymbol{x}$ 称为**豪斯霍尔德反射**. 这样的反射在计算机程序中用来在一个向量（通常是矩阵的列）中产生更多的零.

14. 设 $\boldsymbol{u}=\begin{bmatrix}0\\0\\1\end{bmatrix}$ 和 $\boldsymbol{x}=\begin{bmatrix}1\\5\\3\end{bmatrix}$，确定如习题 13 中的 P 和 Q 并计算 $P\boldsymbol{x}$ 和 $Q\boldsymbol{x}$，图 2-29 说明 $Q\boldsymbol{x}$ 是 $\boldsymbol{x}$ 关于 x_1x_2 平面的反射.

15. 设 $C=E_3E_2E_1B$，其中 E_1,E_2,E_3 是初等矩阵，说明为什么 C 行等价于 B.

16. 设 A 是 $n\times n$ 奇异矩阵，如何构造一个 $n\times n$ 非零矩阵 B 使 $AB=0$？

17. 设 A 是一 6×4 矩阵，B 是一 4×6 矩阵，证明 6×6 矩阵 AB 不可逆.

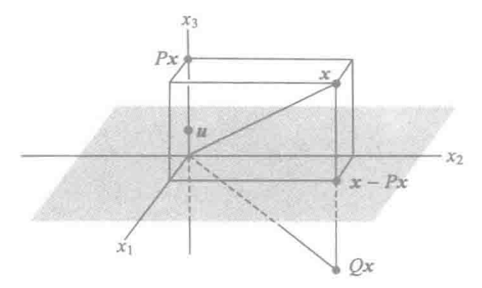


图 2-29　关于平面 $x_3=0$ 的豪斯霍尔德反射

18. 设 A 是 5×3 矩阵，存在 3×5 矩阵 C 使得 $CA=I_3$，又设对 $\mathbb{R}^5$ 中某个 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 有至少一个解，证明方程的解是唯一的.

19. [M]某些动力系统可借助矩阵的幂来研究，如下所示. 给定下列矩阵 A 与 B，确定当 k 增加时（例如 $k=2,\cdots,16$），A^k 与 B^k 有何变化，识别 A 和 B 有什么特点？研究类似矩阵的幂，提出关于这类矩阵的猜想.

$$A=\begin{bmatrix}0.4&0.2&0.3\\0.3&0.6&0.3\\0.3&0.2&0.4\end{bmatrix},\ B=\begin{bmatrix}0&0.2&0.3\\0.1&0.6&0.3\\0.9&0.2&0.4\end{bmatrix}$$

20. [M]设 A_n 为 $n\times n$ 矩阵，主对角线各元素为 0 而其他元素为 1，对 $n=4,5,6$ 计算 A_n^{-1}，对值更大的 n，提出关于一般的 A_n^{-1} 的猜想.

第3章 行 列 式

介绍性实例 随机过程和畸变

诺贝尔物理奖获得者理查德·费曼在他的自传《别闹了，费曼先生》中，述说了他曾在普林斯顿研究生公寓观察蚂蚁一事.他在悬挂的绳子上放置一个纸船通过轮渡来研究蚂蚁的寻食行为，蚂蚁很难找到涂有食饵糖的地方. 然而，当一个蚂蚁爬上纸船后，费曼就将蚂蚁航运到有食饵糖的地方，等蚂蚁获取食物往回搬并爬上纸船后，费曼又将蚂蚁运回到蚂蚁当初上船的地方. 如此以同样方式反复多次，等蚂蚁学会了轮渡之后，费曼更改了轮渡返航的登陆地点，结果使得这群蚂蚁在返回时出现混乱.这说明它们的“学习”是由创造踪迹和跟随踪迹两部分构成的. 费曼又通过地上的载玻片证明了这个猜想. 一旦蚂蚁在载玻片上留下了踪迹，他就在载玻片上重新设置，即重新设置了蚂蚁踪迹的路径，果然，蚂蚁沿着新的踪迹移动，也就是说，蚂蚁是沿着费曼希望的路线移动的.

费曼又用丝网组成的球面来做进一步的研究. 假设一个蚂蚁在球面上沿着独立的路线行走并且在每个交叉口处都要做出选择，即向左走还是向右走. 如果把一些蚂蚁放在球面上，并把同蚂蚁数量具有相等份数的食饵放置在球面上各处，每一只蚂蚁是独立地找到属于自己的那份食物？还是沿着其他蚂蚁的踪迹找到其他蚂蚁的食物？哪种可能性更大?[㊀]

为了记录下蚂蚁的实际行踪，也为了和其他人进行交流，我们可以用多种方法很容易地把球面分割成许多小块. 一种简单的方法是仿照地图用经度和纬度分别作为坐标 x 和 y，虽然这并不能真实地反映整个球面. 赤道附近处两者比较一致，两极附近的扭曲就很大了. 两极区域的像比赤道附近类似区域的像大很多. 为了与地图的这种不同区域的差别相应，极点附近的一只蚂蚁的像应该比赤道附近的像大，那么问题是，大多少呢？

令人惊奇的是，蚂蚁路径和区域扭曲问题都可以通过计算行列式得到很好的解释，行列式正是本章的主题. 行列式有很多应用，比如早在 1900 年，行列式的应用就曾被托马斯·米尔（Thomas Muir）的四卷论著所论述过. 随着矩阵理论的广泛应用，研究的重点发生了变化，行列式在当时的许多很重要的应用在今天好像显得无足轻重了，不过至今行列式本身仍扮演着很重要的角色. 除了在 3.1 节中引入行列式外，这一章还将学习两个重要内容：首先是对第 5 章学习起着重要作用的 3.2 节，方阵可逆性的判别标准；其次是 3.3 节，解释行列式是如何测度一个线性变换对图形面积的改变量的. 当局部地应用时，这一意义回答了两极附近的地图的扩张比例问题. 这种思想以雅可比式在多元微积分中

㊀ 关于反路径问题（和其他两种应用）的解可以参阅 Arthur Benjamin 和 Naomi Cameron 于 2005 年 6 月在 *Mathematical Monthly* 发表的文章.

扮演着重要角色.

>>>>>>>>>

3.1 行列式介绍

针对 2.2 节中的结论，我们回顾一个 2×2 矩阵是可逆的，当且仅当它的行列式非零. 为了将这个有用的结果推广到更大的矩阵，我们需要对 $n\times n$ 矩阵的行列式给出一个定义. 我们可以通过观察对一个可逆的 3×3 矩阵 A 做行化简时会出现什么结果来发现 3×3 矩阵的行列式的定义.

考虑 $A=[a_{ij}]$, $a_{11}\neq 0$，A 的第二行和第三行都乘以 a_{11}，然后再分别减去第一行适当的倍数，则 A 行等价于下面两个矩阵：

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{11}a_{21} & a_{11}a_{22} & a_{11}a_{23} \\ a_{11}a_{31} & a_{11}a_{32} & a_{11}a_{33} \end{bmatrix} \sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{11}a_{22}-a_{12}a_{21} & a_{11}a_{23}-a_{13}a_{21} \\ 0 & a_{11}a_{32}-a_{12}a_{31} & a_{11}a_{33}-a_{13}a_{31} \end{bmatrix} \qquad (1)$$

由于 A 可逆，(1) 式右边矩阵中 (2,2) 和 (3,2) 位置的值不同时为 0. 不妨假设 (2,2) 位置的值不等于零（否则，可以做一个行交换变成这种情形）. 第三行乘以 $a_{11}a_{22}-a_{12}a_{21}$，再对这个新的第三行加上第二行乘以 $-(a_{11}a_{32}-a_{12}a_{31})$，这样就有

$$A \sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{11}a_{22}-a_{12}a_{21} & a_{11}a_{23}-a_{13}a_{21} \\ 0 & 0 & a_{11}\Delta \end{bmatrix}$$

这里

$$\Delta = a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31} \qquad (2)$$

由于 A 可逆，Δ 一定不等于零. 反之亦然，我们将在 3.2 节中看到这个结论. 我们称 (2) 式中的 Δ 为 3×3 矩阵 A 的**行列式**.

回忆 2×2 矩阵 $A=[a_{ij}]$ 的行列式，即 $\det A=a_{11}a_{22}-a_{12}a_{21}$. 对 1×1 矩阵，如 $A=[a_{11}]$，定义 $\det A=a_{11}$. 为了将行列式的定义推广到更大的矩阵，我们将利用 2×2 行列式来重写上面叙述的 3×3 行列式 Δ. 因为 Δ 中的项可以写成

$$(a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32})-(a_{12}a_{21}a_{33}-a_{12}a_{23}a_{31})+(a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31})$$

$$\Delta = a_{11}\cdot\det\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12}\cdot\det\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13}\cdot\det\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

为了简单，可写成

$$\Delta = a_{11}\cdot\det A_{11} - a_{12}\cdot\det A_{12} + a_{13}\cdot\det A_{13} \qquad (3)$$

这里 A_{11}，A_{12} 和 A_{13} 由 A 中划去第一行和三列中之一列而得到. 对任意方阵 A，令 A_{ij} 表示通过划掉 A 中第 i 行和第 j 列而得到的子矩阵，比如，若

$$A = \begin{bmatrix} 1 & -2 & 5 & 0 \\ 2 & 0 & 4 & -1 \\ 3 & 1 & 0 & 7 \\ 0 & 4 & -2 & 0 \end{bmatrix}$$

则 A_{32} 通过划掉第三行和第二列而得到

$$\begin{bmatrix} 1 & -2 & 5 & 0 \\ 2 & 0 & 4 & -1 \\ 3 & 1 & 0 & 7 \\ 0 & 4 & -2 & 0 \end{bmatrix}$$

即

$$A_{32} = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

现在我们可以给出行列式的一个递归定义. 当 $n=3$ 时，像上面的（3）式，$\det A$ 由 2×2 子矩阵 A_{1j} 的行列式来定义. 当 $n=4$ 时， $\det A$ 由 3×3 子矩阵 A_{1j} 来定义. 一般情形，一个 $n\times n$ 行列式由 $(n-1)\times(n-1)$ 子矩阵的行列式来定义.

定义 当 $n\geqslant 2$，$n\times n$ 矩阵 $A=[a_{ij}]$ 的行列式是形如 $\pm a_{1j}\det A_{1j}$ 的 n 个项的和，其中加号和减号交替出现，这里元素 $a_{11}, a_{12}, \cdots a_{1n}$ 来自 A 的第一行，即

$$\begin{aligned}\det A &= a_{11}\cdot\det A_{11} - a_{12}\cdot\det A_{12} + \cdots + (-1)^{1+n} a_{1n}\cdot\det A_{1n} \\ &= \sum_{j=1}^{n}(-1)^{1+j} a_{1j}\det A_{1j}\end{aligned}$$

例 1 计算行列式 $\det A$，其中

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

解 计算 $\det A = a_{11}\cdot\det A_{11} - a_{12}\cdot\det A_{12} + a_{13}\cdot\det A_{13}$.

$$\begin{aligned}\det A &= 1\cdot\det\begin{bmatrix} 4 & -1 \\ -2 & 0 \end{bmatrix} - 5\cdot\det\begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} + 0\cdot\det\begin{bmatrix} 2 & 4 \\ 0 & -2 \end{bmatrix} \\ &= 1(0-2) - 5(0-0) + 0(-4-0) = -2\end{aligned}$$

■

方阵的行列式的另一个常用记号是利用一对竖线代替括号. 这样，例 1 中的运算可以写成

$$\det A = 1\cdot\begin{vmatrix} 4 & -1 \\ -2 & 0 \end{vmatrix} - 5\cdot\begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix} + 0\cdot\begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix} = \cdots = -2$$

为了叙述下一个定理，将 $\det A$ 的定义用稍微不同的形式写出是很方便的. 给定 $A=[a_{ij}]$，A 的 (i, j) **余因子** C_{ij} 由下式给出

$$C_{ij} = (-1)^{i+j}\det A_{ij} \tag{4}$$

则

$$\det A = a_{11}\cdot C_{11} + a_{12}\cdot C_{12} + \cdots + a_{1n}\cdot C_{1n}$$

这个公式称为按 A 的**第一行的余因子展开式**，为了避免冗长的陈述，我们省略下面基本定理的证明.

定理 1 $n\times n$ 矩阵 A 的行列式可按任意行或列的余因子展开式来计算. 按第 i 行展开用(4)

式给出的余因子写法可以写成：

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

按第 j 列的余因子展开式为

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

(i,j) 位置的余因子中加号或减号取决于 a_{ij} 在矩阵中的位置，而与 a_{ij} 本身的符号无关，因子 $(-1)^{i+j}$ 可确定下面符号的棋盘模式：

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \\ + & - & + & \\ \vdots & & & \ddots \end{bmatrix}$$

例 2 利用按第三行的余因子展开式求 $\det A$，其中

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

解 计算

$$\begin{aligned} \det A &= a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33} \\ &= (-1)^{3+1}a_{31}\det A_{31} + (-1)^{3+2}a_{32}\det A_{32} + (-1)^{3+3}a_{33}\det A_{33} \\ &= 0\begin{vmatrix} 5 & 0 \\ 4 & -1 \end{vmatrix} - (-2)\begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} + 0\begin{vmatrix} 1 & 5 \\ 2 & 4 \end{vmatrix} = 0 + 2(-1) + 0 = -2 \end{aligned}$$

■

定理 1 有助于计算包含有许多零的行列式. 例如，如果某行多数元素为零，则按此行的余因子展开式就会有许多项为零，这些项中的余因子就不需计算. 同理对多数元素为零的某列也是这样.

例 3 计算 $\det A$，这里

$$A = \begin{bmatrix} 3 & -7 & 8 & 9 & -6 \\ 0 & 2 & -5 & 7 & 3 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 2 & 4 & -1 \\ 0 & 0 & 0 & -2 & 0 \end{bmatrix}$$

解 按 A 的第一列的余因子展开式中，除第一项外均为零，从而

$$\det A = 3\cdot\begin{bmatrix} 2 & -5 & 7 & 3 \\ 0 & 1 & 5 & 0 \\ 0 & 2 & 4 & -1 \\ 0 & 0 & -2 & 0 \end{bmatrix} - 0\cdot C_{21} + 0\cdot C_{31} - 0\cdot C_{41} + 0\cdot C_{51}$$

接下来在余因子展开式中省略零项，再利用第一列零的优势，按第一列展开这个 4×4 行列式，有

$$\det A = 3\cdot 2\cdot\begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

这个 3×3 行列式已在例 1 中算出结果为 –2，从而 $\det A = 3\cdot 2\cdot(-2) = -12$ ■

例 3 中矩阵接近三角阵，此例中的方法容易用来证明下面的定理.

定理 2 若 A 为三角阵，则 $\det A$ 等于 A 的主对角线上元素的乘积.

当某行或某列全为零时，例 3 中找零的策略效果特别好，在这种情形下，按这一行或列的余因子展开式是一些零的和，从而此行列式为零. 但不幸的是，大多数的余因子展开式并不是这样快能求值的.

数值计算的注解 按现在的标准来说，一个 25×25 矩阵还是小的. 但是用余因子展开式来计算一个 25×25 行列式是不可行的. 一般对 $n\times n$ 行列式而言，余因子展开式需要计算超过 $n!$ 个乘法运算，25! 近似于 1.5×10^{25}.

若一个超级计算机每秒钟能够完成 1 万亿次乘法运算，用这种方法计算一个 25×25 行列式，它将需运行 50 万年. 幸运的是，稍候我们就会发现快速计算的方法.

在习题 19~38 中，多数是 2×2 情形的矩阵，揭示了行列式的一些重要性质. 习题 33~36 的结果将用来在下一节导出 $n\times n$ 矩阵的与之相似的性质.

练习题

计算 $\begin{vmatrix} 5 & -7 & 2 & 2 \\ 0 & 3 & 0 & -4 \\ -5 & -8 & 0 & 3 \\ 0 & 5 & 0 & -6 \end{vmatrix}$.

习题 3.1

利用第一行的余因子展开式计算下列 1~8 题中的行列式. 按第二列的余因子展开式再次计算 1~4 题中的行列式.

1. $\begin{vmatrix} 3 & 0 & 4 \\ 2 & 3 & 2 \\ 0 & 5 & -1 \end{vmatrix}$

2. $\begin{vmatrix} 0 & 5 & 1 \\ 4 & -3 & 0 \\ 2 & 4 & 1 \end{vmatrix}$

3. $\begin{vmatrix} 2 & -4 & 3 \\ 3 & 1 & 2 \\ 1 & 4 & -1 \end{vmatrix}$

4. $\begin{vmatrix} 1 & 3 & 5 \\ 2 & 1 & 1 \\ 3 & 4 & 2 \end{vmatrix}$

5. $\begin{vmatrix} 2 & 3 & -4 \\ 4 & 0 & 5 \\ 5 & 1 & 6 \end{vmatrix}$

6. $\begin{vmatrix} 5 & -2 & 4 \\ 0 & 3 & -5 \\ 2 & -4 & 7 \end{vmatrix}$

7. $\begin{vmatrix} 4 & 3 & 0 \\ 6 & 5 & 2 \\ 9 & 7 & 3 \end{vmatrix}$

8. $\begin{vmatrix} 8 & 1 & 6 \\ 4 & 0 & 3 \\ 3 & -2 & 5 \end{vmatrix}$

利用余因子展开式计算 9~14 题中的行列式，在每一步，选择使得计算量最少的行和列.

9. $\begin{vmatrix} 6 & 0 & 0 & 5 \\ 1 & 7 & 2 & -5 \\ 2 & 0 & 0 & 0 \\ 8 & 3 & 1 & 8 \end{vmatrix}$

10. $\begin{vmatrix} 1 & -2 & 5 & 2 \\ 0 & 0 & 3 & 0 \\ 2 & -6 & -7 & 5 \\ 5 & 0 & 4 & 4 \end{vmatrix}$

11. $\begin{vmatrix} 3 & 5 & -8 & 4 \\ 0 & -2 & 3 & -7 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 2 \end{vmatrix}$

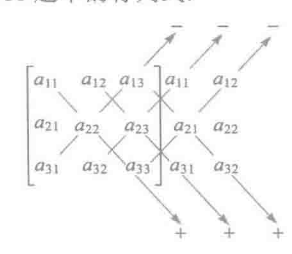
12. $\begin{vmatrix} 4 & 0 & 0 & 0 \\ 7 & -1 & 0 & 0 \\ 2 & 6 & 3 & 0 \\ 5 & -8 & 4 & -3 \end{vmatrix}$

13. $\begin{vmatrix} 4 & 0 & -7 & 3 & -5 \\ 0 & 0 & 2 & 0 & 0 \\ 7 & 3 & -6 & 4 & -8 \\ 5 & 0 & 5 & 2 & -3 \\ 0 & 0 & 9 & -1 & 2 \end{vmatrix}$

14. $\begin{vmatrix} 6 & 3 & 2 & 4 & 0 \\ 9 & 0 & -4 & 1 & 0 \\ 8 & -5 & 6 & 7 & 1 \\ 3 & 0 & 0 & 0 & 0 \\ 4 & 2 & 3 & 2 & 0 \end{vmatrix}$

一个 3×3 行列式的展开式可由下列方法记忆，在矩阵的右边分别再写出矩阵的前两列，通过先算

出六个对角线上数字的乘积，然后箭头向下的三个对角线上数字的乘积相加再减去箭头向上的三个对角线上数字的乘积就算出了这个行列式. 用此法计算 15~18 题中的行列式.



注意，这种方法不能推广到 4×4 或更大的矩阵.

15. $\begin{vmatrix} 3 & 0 & 4 \\ 2 & 3 & 2 \\ 0 & 5 & -1 \end{vmatrix}$　16. $\begin{vmatrix} 0 & 5 & 1 \\ 4 & -3 & 0 \\ 2 & 4 & 1 \end{vmatrix}$

17. $\begin{vmatrix} 2 & -4 & 3 \\ 3 & 1 & 2 \\ 1 & 4 & -1 \end{vmatrix}$　18. $\begin{vmatrix} 1 & 3 & 5 \\ 2 & 1 & 1 \\ 3 & 4 & 2 \end{vmatrix}$

在习题 19~24 中，考察矩阵的行列式在初等行变换下的变化. 对每一个情形，说出是何种行变换，再说出此变换对这个行列式有何影响.

19. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} c & d \\ a & b \end{bmatrix}$　20. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} a & b \\ kc & kd \end{bmatrix}$

21. $\begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}, \begin{bmatrix} 3 & 4 \\ 5+3k & 6+4k \end{bmatrix}$

22. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} a+kc & b+kd \\ c & d \end{bmatrix}$

23. $\begin{bmatrix} 1 & 1 & 1 \\ -3 & 8 & -4 \\ 2 & -3 & 2 \end{bmatrix}, \begin{bmatrix} k & k & k \\ -3 & 8 & -4 \\ 2 & -3 & 2 \end{bmatrix}$

24. $\begin{bmatrix} a & b & c \\ 3 & 2 & 2 \\ 6 & 5 & 6 \end{bmatrix}, \begin{bmatrix} 3 & 2 & 2 \\ a & b & c \\ 6 & 5 & 6 \end{bmatrix}$

计算习题 25~30 给出的初等矩阵的行列式(见 2.2 节).

25. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & k & 1 \end{bmatrix}$　26. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{bmatrix}$

27. $\begin{bmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$　28. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}$

29. $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$　30. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

利用 25~28 题回答习题 31 和习题 32 中的问题，给出答案的理由.

31. 初等矩阵经过行替代（即行倍加），其行列式是什么？
32. 对角线上某元素乘 k 的初等矩阵的行列式是什么？

在 33~36 题中，证明 $\det EA = (\det E)(\det A)$. 这里 E 为给出的初等矩阵，$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

33. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$　34. $\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$

35. $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$　36. $\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$

37. $A = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix}$，写出 $5A$，$\det 5A = 5\det A$ 吗？
38. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$，$k$ 为一个数，找出 $\det kA$ 与 k 和 $\det A$ 相联系的公式.

在习题 39 和习题 40 中，A 为 $n\times n$ 矩阵，标出每个命题的真假，验证每个回答.

39. a. 一个 $n\times n$ 行列式由 $(n-1)\times(n-1)$ 子矩阵的行列式所定义.
    b. 矩阵 A 的 (i, j) 余因子是 A 中划掉第 i 行和第 j 列所得到的矩阵 A_{ij}.
40. a. $\det A$ 的按某列的余因子展开式是按某行展开式的−1 倍.
    b. 三角形矩阵的行列式等于主对角线上元素之和.
41. 若 $\boldsymbol{u} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$，$\boldsymbol{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$，计算由 $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{u}+\boldsymbol{v}$ 和 $\boldsymbol{0}$ 构成的平行四边形的面积，再计算 $[\boldsymbol{u}\ \ \boldsymbol{v}]$ 的行列

式，二者相比较，结果如何. 将 v 的第一个数字换成任意数 x，重复考虑上面的问题，画一个图并解释你的发现.

42. 令 $\boldsymbol{u}=\begin{bmatrix}a\\b\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}c\\0\end{bmatrix}$，这里 a,b,c 是正数（为了简单起见），计算由 $\boldsymbol{u},\boldsymbol{v},\boldsymbol{u}+\boldsymbol{v}$ 和 $\boldsymbol{0}$ 构成的平行四边形的面积，再计算 $[\boldsymbol{u}\ \ \boldsymbol{v}]$ 和 $[\boldsymbol{v}\ \ \boldsymbol{u}]$ 的行列式，画一个图并解释你的发现.

43. [M] $\det(A+B)=\det A+\det B$ 对吗？为看到结果是否正确，随意给两个 5×5 矩阵 A,B，计算 $\det(A+B)-\det A-\det B$（参考 2.1 节中习题 37）. 对另外 3 对不同的 $n\times n$ 矩阵，重复上述计算，报告你的结论.

44. [M] $\det AB=(\det A)(\det B)$ 对吗？像 43 题一样用 4 对随机矩阵来检验，并给出一个猜想.

45. [M]随意构造一个元素均为从−9 到 9 的整数的 4×4 矩阵，计算 $\det A$，$\det A^{\mathrm{T}}$，$\det(-A)$，$\det(2A)$ 和 $\det(10A)$. 对另外两个随机的 4×4 整数矩阵重复以上计算，给出这些行列式之间关系的一个猜想.（参考 2.1 节习题 36.）用几个随机的 5×5 和 6×6 整数矩阵来检验你的猜想，若有必要，修正你的猜想，报告你的结果.

46. [M] $\det A^{-1}$ 与 $\det A$ 关系如何？用随机的 $n\times n$ 整数矩阵来检验 $(n=4,5,6)$. 并做出猜想.

注：万一你遇到 $\det A=0$ 的情形，化简矩阵 A 成阶梯形的情形，再讨论你的发现.

练习题答案

充分利用 0，开始先按第 3 列的余因子展开成为 3 阶行列式，对此再按第一列展开：

$$\begin{vmatrix}5&-7&2&2\\0&3&0&-4\\-5&-8&0&3\\0&5&0&-6\end{vmatrix}=(-1)^{1+3}\cdot2\begin{vmatrix}0&3&-4\\-5&-8&3\\0&5&-6\end{vmatrix}=2\cdot(-1)^{2+1}(-5)\begin{vmatrix}3&-4\\5&-6\end{vmatrix}=20$$

计算中倒数第 2 步中的 $(-1)^{2+1}$ 是由 3×3 行列式中−5 的(2,1)位置决定的.

3.2 行列式的性质

行列式的奥秘在于当做行变换时它如何变化. 下列定理推广了 3.1 节中习题 19~24 的结果，其证明在本节末尾.

定理 3 （行变换）

令 A 是一个方阵.

a. 若 A 的某一行的倍数加到另一行得矩阵 B，则 $\det B=\det A$.

b. 若 A 的两行互换得矩阵 B，则 $\det B=-\det A$.

c. 若 A 的某行乘以 k 倍得到矩阵 B，则 $\det B=k\cdot\det A$.

下列例子展示如何利用定理 3 来有效地计算行列式.

例 1 计算 $\det A$，其中 $A=\begin{bmatrix}1&-4&2\\-2&8&-9\\-1&7&0\end{bmatrix}$.

解 思路是先将 A 化简成阶梯形，再利用三角形矩阵的行列式等于对角线上元素之积的知

识. 针对第一列的前两次行倍加均不改变行列式的值.

$$\det A=\begin{vmatrix}1&-4&2\\-2&8&-9\\-1&7&0\end{vmatrix}=\begin{vmatrix}1&-4&2\\0&0&-5\\-1&7&0\end{vmatrix}=\begin{vmatrix}1&-4&2\\0&0&-5\\0&3&2\end{vmatrix}$$

交换第 2 行与第 3 行使行列式取反号，即

$$\det A=-\begin{vmatrix}1&-4&2\\0&3&2\\0&0&-5\end{vmatrix}=-(1)(3)(-5)=15$$

■

计算行列式时，通常利用定理 3(c)将某一行的公因子提出来，如：

$$\begin{vmatrix}*&*&*\\5k&-2k&3k*&*&*\end{vmatrix}=k\begin{vmatrix}*&*&*\\5&-2&3*&*&*\end{vmatrix}$$

这里用星号标记的元素不变. 在下例中用此方法.

例 2 计算 $\det A$，这里 $A=\begin{bmatrix}2&-8&6&8\\3&-9&5&10\\-3&0&1&-2\\1&-4&0&6\end{bmatrix}$.

解 为简化计算，设法使左上角为 1，可将第 1 行与第 4 行交换，也可由第 1 行提出因子 2，再对第 1 列做行倍加：

$$\det A=2\begin{vmatrix}1&-4&3&4\\3&-9&5&10\\-3&0&1&-2\\1&-4&0&6\end{vmatrix}=2\begin{vmatrix}1&-4&3&4\\0&3&-4&-2\\0&-12&10&10\\0&0&-3&2\end{vmatrix}$$

然后，可再从第 3 行中提出 2，或利用第 2 列中的 3 作为一个主元. 此处，我们用后者，将第 2 行的 4 倍加到第 3 行：

$$\det A=2\begin{vmatrix}1&-4&3&4\\0&3&-4&-2\\0&0&-6&2\\0&0&-3&2\end{vmatrix}$$

最后，将 $-1/2$ 倍的第 3 行加到第 4 行，再计算这个“三角形”行列式得：

$$\det A=2\begin{vmatrix}1&-4&3&4\\0&3&-4&-2\\0&0&-6&2\\0&0&0&1\end{vmatrix}=2\cdot(1)\cdot(3)\cdot(-6)\cdot(1)=-36$$

■

若一个方阵 A 被行倍加和行交换化简为阶梯形 U（这总是可以做到的，见 1.2 节行化简算法），且此过程经过了 r 次行交换，则定理 3 表明

$$\det A = (-1)^r \det U$$

由于U是阶梯形，它是三角阵，因此$\det U$是主对角线上的元素$u_{11},\cdots,u_{nn}$的乘积.若A可逆，则元素u_{ii}都是主元（因为$A \sim I_n$且u_{ii}没有被倍乘变为1）. 否则，至少有u_{nn}等于零，乘积$u_{11}\cdots u_{nn}$为零. 见图3-1.

$$U=\begin{bmatrix} ■ & * & * & * \\ 0 & ■ & * & * \\ 0 & 0 & ■ & * \\ 0 & 0 & 0 & ■ \end{bmatrix}$$

$\det U \neq 0$

$$U=\begin{bmatrix} ■ & * & * & * \\ 0 & ■ & * & * \\ 0 & 0 & 0 & ■ \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$\det U = 0$

图3-1 标准的阶梯形方阵

从而有以下公式

$$\det A=\begin{cases}(-1)^r \cdot (U\text{的主元乘积}) & \text{当}A\text{可逆} \\ 0 & \text{当}A\text{不可逆}\end{cases} \tag{1}$$

注意到下列有趣的情形：尽管上述中的阶梯形U是不唯一的（因为它并没有经过完全的行简化），主元也不是唯一的，但除了差一个负号外，这些主元的乘积是唯一的.

公式（1）不但给出$\det A$的一个具体的解释，还证明了本节的主要定理：

定理4 方阵A是可逆的当且仅当$\det A \neq 0$.

定理4把语句“$\det A \neq 0$”增加到可逆矩阵定理中. 一个有用的推论是若A的列是线形相关的，则$\det A = 0$. 而且若A的行是线形相关的，则$\det A = 0$.（A的行是A^{T}的列，由A^{T}的列线性相关可推出A^{T}是奇异的. 当A^{T}是奇异矩阵时，由可逆矩阵定理可知，A也是奇异的.）在实际问题中，当两行或两列是相同的或者一行或一列是零时，则线形相关是显然的.

例3 计算$\det A$，$A=\begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ -6 & 7 & -7 & 4 \\ -5 & -8 & 0 & 9 \end{bmatrix}$.

解 将2倍的第1行加到第3行，得

$$\det A = \det\begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ 0 & 5 & -3 & -6 \\ -5 & -8 & 0 & 9 \end{bmatrix} = 0$$

这是因为第二个矩阵的第 2 行和第 3 行相等. ■

数值计算的注解

1. 对一般的矩阵 A，许多计算 $\det A$ 的计算机程序使用上面公式（1）的方法.

2. 可以证明一个 $n\times n$ 行列式利用行变换展开大约需要 $2n^3/3$ 次算术运算，任何现代微型计算机可以在不到一秒钟内计算一个 25×25 行列式，因为仅需大约 10 000 次运算.

充分利用出现的多个零，利用特别的指令，计算机还可以处理大的“稀疏”矩阵. 当然，零元素可加快手算过程. 下例的计算中结合了行变换的优势和 3.1 节中余因子展开中利用零元素的技巧.

例 4 计算 $\det A$，$A=\begin{bmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ -2 & -5 & 4 & -2 \end{bmatrix}$.

解 开始最好是利用第 1 列中的 2 作为一个主元，消去下面的-2，再利用余因子展开简化成低一阶的行列式，然后再由行倍加变换，有

$$\det A=\begin{vmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ 0 & 0 & -3 & 1 \end{vmatrix}=-2\begin{vmatrix} 1 & 2 & -1 \\ 3 & 6 & 2 \\ 0 & -3 & 1 \end{vmatrix}=-2\begin{vmatrix} 1 & 2 & -1 \\ 0 & 0 & 5 \\ 0 & -3 & 1 \end{vmatrix}$$

交换第 2 行与第 3 行得到一个“三角形行列式”. 另一种方法是按第 1 列的余因子展开：

$$\det A=(-2)\cdot(1)\begin{bmatrix} 0 & 5 \\ -3 & 1 \end{bmatrix}=-2\cdot(15)=-30$$ ■

列变换

类似于前面已经考虑过的行变换，我们可以对矩阵的列实行变换，下一个定理证明行列式的列变换与行变换具有相同的效果.

定理 5 若 A 为一个 $n\times n$ 矩阵，则 $\det A^{\mathrm{T}}=\det A$.

证明 当 $n=1$ 时，定理显然成立.假设定理对 $k\times k$ 行列式成立. 令 $n=k+1$，则 A 中 a_{1j} 的余因子等于 A^{T} 中 a_{j1} 的余因子，因为这些余因子是由 $k\times k$ 行列式表示的，所以 $\det A$ 沿着第一行的余因子展开等于 $\det A^{\mathrm{T}}$ 沿第一列的余因子展开，即 $\det A=\det A^{\mathrm{T}}$. 于是定理对 $n=k+1$ 也成立，即定理对一个 n 成立就可以推出对 $n+1$ 成立. 由归纳法原理，定理对任意 $n\geqslant 1$ 均成立. ■

因为定理 5 成立，所以当把定理 3 中出现的“行”字变成“列”字时，定理 3 中每一个命题均成立. 为证此性质，只需针对 A^{T} 应用原来的定理 3. 对 A^{T} 的一个行变换相当于对 A 的一个列变换.

列变换对理论和手工计算都是很有用的. 然而，为了简单，在数值计算中，我们仅实行行变换.

行列式与矩阵乘积

下列有用定理的证明放在本节末尾，其应用放在习题中.

定理 6 （乘法的性质）
若 A 和 B 均为 $n\times n$ 矩阵，则 $\det AB = (\det A)(\det B)$.

例 5 对 $A=\begin{bmatrix}6 & 1\\ 3 & 2\end{bmatrix}$, $B=\begin{bmatrix}4 & 3\\ 1 & 2\end{bmatrix}$，验证定理 6.

解 $AB=\begin{bmatrix}6 & 1\\ 3 & 2\end{bmatrix}\begin{bmatrix}4 & 3\\ 1 & 2\end{bmatrix}=\begin{bmatrix}25 & 20\\ 14 & 13\end{bmatrix}$

$$\det AB = 25\cdot 13-20\cdot 14=325-280=45$$

由于 $\det A=9$ ， $\det B=5$ ，

$$(\det A)(\det B)=9\cdot 5=45=\det AB$$ ■

警告 通常一个错误的观点是定理 6 对矩阵的和也有类似的结果. 然而，一般而言，$\det(A+B)$ 不等于 $\det A+\det B$.

行列式函数的一个线性性质

若 A 为 $n\times n$ 矩阵，我们可以将 $\det A$ 看作 A 中 n 个列向量的函数，我们将证明如果 A 中除了一列之外都是固定的向量，则 $\det A$ 是那个可变列向量的线性函数.

假设 A 的第 j 列允许有变化，A 可写成：

$$A=[\boldsymbol{a}_1 \ \cdots \ \boldsymbol{a}_{j-1} \ \ \boldsymbol{x} \ \ \boldsymbol{a}_{j+1}\cdots \ \boldsymbol{a}_n]$$

定义由 $\mathbb{R}^n$ 到 $\mathbb{R}$ 的变换 T 为：

$$T(\boldsymbol{x})=\det[\boldsymbol{a}_1 \ \cdots \ \boldsymbol{a}_{j-1} \ \ \boldsymbol{x} \ \ \boldsymbol{a}_{j+1} \ \cdots \ \boldsymbol{a}_n]$$

则有

$$T(c\boldsymbol{x})=cT(\boldsymbol{x})\text{，对任意常数 } c \text{ 和 } \mathbb{R}^n \text{ 中任意 } \boldsymbol{x} \text{ 成立} \tag{2}$$

$$T(\boldsymbol{u}+\boldsymbol{v})=T(\boldsymbol{u})+T(\boldsymbol{v})\text{，对 } \mathbb{R}^n \text{ 中所有 } \boldsymbol{u},\boldsymbol{v} \text{ 成立} \tag{3}$$

性质（2）即定理 3(c)应用到 A 的列. （3）的一个证明可由 $\det A$ 按第 j 列的余因子展开来完成（见习题 43）. 行列式的这个线性性质有许多有用的推论，将会在高级的课程学习中研究.

定理 3 和定理 6 的证明

借助于 2.2 节中讨论的初等矩阵证明定理 3 是方便的. 我们称一个初等矩阵 E 是一个行倍加(矩阵)，如果 E 是由单位矩阵 I 经一行加另一行的倍数而得到；称 E 是一个交换，如果 E 是由交换 I 的两行而得到；称 E 是一个 r 倍乘，如果 E 是由 I 的某一行乘以一个非零数量而得到. 用这些术语，定理 3 可重新叙述如下：

若 A 是一个 $n\times n$ 矩阵，E 是一个 $n\times n$ 初等矩阵，则

$$\det EA = (\det E)(\det A)$$

这里

$$\det E = \begin{cases} 1 & 若E是一个行倍加 \\ -1 & 若E是一个交换 \\ r & 若E是一个r倍乘 \end{cases}$$

定理 3 的证明 针对 A 的行或列的大小用归纳法，2×2 矩阵的情形在 3.1 节中习题 33~36 已证. 假设该定理对 $k\times k$ 矩阵的行列式成立，$k\geqslant 2$. 令 $n=k+1$，A 为 $n\times n$ 矩阵. E 对 A 的作用涉及两行或一行. 所以我们可以按在 E 的作用下没有被改变的一行展开 $\det EA$，比如说第 i 行. 令 A_{ij}（分别地，B_{ij}）是由 A（分别地，EA）中划掉第 i 行第 j 列得到的矩阵，则 B_{ij} 的行由 A_{ij} 的行通过实行与 E 作用 A 相同类型的初等行变换得到，由于这些子矩阵仅仅是 $k\times k$ 矩阵，归纳假设蕴涵

$$\det B_{ij} = \alpha\cdot\det A_{ij}$$

这里 $\alpha=1,\ -1,r$，依 E 的类型而定. $\det EA$ 沿第 i 行的余因子展开式是

$$\begin{aligned}\det EA &= a_{i1}(-1)^{i+1}\det B_{i1}+\cdots+a_{in}(-1)^{i+n}\det B_{in}\\ &=\alpha\ a_{i1}(-1)^{i+1}\det A_{i1}+\cdots+\alpha\ a_{in}(-1)^{i+n}\det A_{in}\\ &=\alpha\cdot\det A\end{aligned}$$

特别地，取 $A=I_n$，则 $\det E=1,-1,r$，依赖于 E 的类型而定. 于是定理对 $n=2$ 成立. 定理对 n 成立蕴涵对 $n+1$ 时也成立. 由归纳法原理，定理对 $n\geqslant 2$ 均成立. 对 $n=1$ 时，定理显然成立. ■

定理 6 的证明 若 A 不可逆，由 2.3 节中习题 27，则 AB 也不可逆. 在此情形下，由定理 4，因 $\det A$ 和 $\det A\cdot\det B$ 均为零，则 $\det AB=(\det A)(\det B)$ 成立. 若 A 可逆，由可逆矩阵定理，A 与单位矩阵 I_n 行等价，所以存在初等矩阵 $E_1,\cdots,E_p$ 使得

$$A=E_pE_{p-1}\cdots E_1I_n=E_pE_{p-1}\cdots E_1$$

为了简单，用 $|A|$ 表示 $\det A$. 反复应用定理 3，如上面所重新描述的那样，得

$$\begin{aligned}|AB| &= |E_p\cdots E_1B| = |E_p|\cdot|E_{p-1}\cdots E_1B| = \cdots\\ &= |E_p|\cdots|E_1|\cdot|B| = \cdots = |E_p\cdots E_1|\cdot|B|\\ &= |A|\ |B|\end{aligned}$$

■

练习题

1. 用尽可能少的步骤计算 $\begin{vmatrix} 1 & -3 & 1 & -2 \\ 2 & -5 & -1 & -2 \\ 0 & -4 & 5 & 1 \\ -3 & 10 & -6 & 8 \end{vmatrix}$.

2. 用一个行列式来确定向量 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ 是否线性无关，这里

$$\boldsymbol{v}_1=\begin{bmatrix}5\\-7\\9\end{bmatrix},\ \boldsymbol{v}_2=\begin{bmatrix}-3\\3\\-5\end{bmatrix},\ \boldsymbol{v}_3=\begin{bmatrix}2\\-7\\5\end{bmatrix}$$

习题 3.2

在习题 1~4 中，每个方程说明行列式的一个性质，叙述这些性质.

1. $\begin{vmatrix} 0 & 5 & -2 \\ 1 & -3 & 6 \\ 4 & -1 & 8 \end{vmatrix} = -\begin{vmatrix} 1 & -3 & 6 \\ 0 & 5 & -2 \\ 4 & -1 & 8 \end{vmatrix}$

2. $\begin{vmatrix} 2 & -6 & 4 \\ 3 & 5 & -2 \\ 1 & 6 & 3 \end{vmatrix} = 2\begin{vmatrix} 1 & -3 & 2 \\ 3 & 5 & -2 \\ 1 & 6 & 3 \end{vmatrix}$

3. $\begin{vmatrix} 1 & 3 & -4 \\ 2 & 0 & -3 \\ 5 & -4 & 7 \end{vmatrix} = \begin{vmatrix} 1 & 3 & -4 \\ 0 & -6 & 5 \\ 5 & -4 & 7 \end{vmatrix}$

4. $\begin{vmatrix} 1 & 2 & 3 \\ 0 & 5 & -4 \\ 3 & 7 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 5 & -4 \\ 0 & 1 & -5 \end{vmatrix}$

在习题 5~10 中，通过行化简成阶梯形求行列式的值.

5. $\begin{vmatrix} 1 & 5 & -6 \\ -1 & -4 & 4 \\ -2 & -7 & 9 \end{vmatrix}$

6. $\begin{vmatrix} 1 & 5 & -3 \\ 3 & -3 & 3 \\ 2 & 13 & -7 \end{vmatrix}$

7. $\begin{vmatrix} 1 & 3 & 0 & 2 \\ -2 & -5 & 7 & 4 \\ 3 & 5 & 2 & 1 \\ 1 & -1 & 2 & -3 \end{vmatrix}$

8. $\begin{vmatrix} 1 & 3 & 3 & -4 \\ 0 & 1 & 2 & -5 \\ 2 & 5 & 4 & -3 \\ -3 & -7 & -5 & 2 \end{vmatrix}$

9. $\begin{vmatrix} 1 & -1 & -3 & 0 \\ 0 & 1 & 5 & 4 \\ -1 & 2 & 8 & 5 \\ 3 & -1 & -2 & 3 \end{vmatrix}$

10. $\begin{vmatrix} 1 & 3 & -1 & 0 & -2 \\ 0 & 2 & -4 & -1 & -6 \\ -2 & -6 & 2 & 3 & 9 \\ 3 & 7 & -3 & 8 & -7 \\ 3 & 5 & 5 & 2 & 7 \end{vmatrix}$

在习题 11~14 中，结合行简化和余因子展开的方法计算行列式的值.

11. $\begin{vmatrix} 2 & 5 & -3 & -1 \\ 3 & 0 & 1 & -3 \\ -6 & 0 & -4 & 9 \\ 4 & 10 & -4 & -1 \end{vmatrix}$

12. $\begin{vmatrix} -1 & 2 & 3 & 0 \\ 3 & 4 & 3 & 0 \\ 5 & 4 & 6 & 6 \\ 4 & 2 & 4 & 3 \end{vmatrix}$

13. $\begin{vmatrix} 2 & 5 & 4 & 1 \\ 4 & 7 & 6 & 2 \\ 6 & -2 & -4 & 0 \\ -6 & 7 & 7 & 0 \end{vmatrix}$

14. $\begin{vmatrix} -3 & -2 & 1 & -4 \\ 1 & 3 & 0 & -3 \\ -3 & 4 & -2 & 8 \\ 3 & -4 & 0 & 4 \end{vmatrix}$

求习题 15~20 中的行列式，已知 $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = 7$.

15. $\begin{vmatrix} a & b & c \\ d & e & f \\ 5g & 5h & 5i \end{vmatrix}$

16. $\begin{vmatrix} a & b & c \\ 3d & 3e & 3f \\ g & h & i \end{vmatrix}$

17. $\begin{vmatrix} a & b & c \\ g & h & i \\ d & e & f \end{vmatrix}$

18. $\begin{vmatrix} g & h & i \\ a & b & c \\ d & e & f \end{vmatrix}$

19. $\begin{vmatrix} a & b & c \\ 2d+a & 2e+b & 2f+c \\ g & h & i \end{vmatrix}$

20. $\begin{vmatrix} a+d & b+e & c+f \\ d & e & f \\ g & h & i \end{vmatrix}$

在习题 21~23 中，利用行列式检验这些矩阵是否可逆.

21. $\begin{bmatrix} 2 & 3 & 0 \\ 1 & 3 & 4 \\ 1 & 2 & 1 \end{bmatrix}$

22. $\begin{bmatrix} 5 & 0 & -1 \\ 1 & -3 & -2 \\ 0 & 5 & 3 \end{bmatrix}$

23. $\begin{bmatrix} 2 & 0 & 0 & 8 \\ 1 & -7 & -5 & 0 \\ 3 & 8 & 6 & 0 \\ 0 & 7 & 5 & 4 \end{bmatrix}$

在习题 24~26 中，利用行列式判断每组向量是否线性无关.

24. $\begin{bmatrix} 4 \\ 6 \\ -7 \end{bmatrix}, \begin{bmatrix} -7 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} -3 \\ -5 \\ 6 \end{bmatrix}$

25. $\begin{bmatrix} 7 \\ -4 \\ -6 \end{bmatrix}, \begin{bmatrix} -8 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 7 \\ 0 \\ -5 \end{bmatrix}$

26. $\begin{bmatrix} 3 \\ 5 \\ -6 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ -6 \\ 0 \\ 7 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ -3 \end{bmatrix}$

在习题 27 和习题 28 中，A,B 均为 $n\times n$ 矩阵，标出每个命题的真假，给出理由.

27. a. 矩阵的行倍加变换不影响矩阵的行列式.
    b. 矩阵 A 的行列式是 A 的任意一个阶梯形 U 中主元之积再乘以 $(-1)^r$. 这里 r 是 A 经行简化成 U 时做的行交换的次数.
    c. 若 A 的列是线性相关的，则 $\det A=0$.
    d. $\det(A+B)=\det A+\det B$.
28. a. 如果连续作了两次行交换，则新的行列式等于原来的行列式.
    b. A 的行列式等于其对角线元素之积.
    c. 若 $\det A$ 为零，则其两行或两列相同，或者，一行或一列为零.
    d. $\det A^{\mathrm{T}}=(-1)\det A$.
29. 计算 $\det B^5$，其中 $B=\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}$.
30. 利用定理 3（但不用定理 4）来证明，若方阵 A 的两行相等，则 $\det A=0$. 此结论对两列相等的情形也成立，为什么？

对习题 31~36，在你的解释中指出用到的一个合适的定理.

31. 证明：如果 A 可逆，则 $\det A^{-1}=\dfrac{1}{\det A}$.
32. 若 A 为一个 $n\times n$ 矩阵，试找出一个计算 $\det(rA)$ 的公式.
33. 若 A,B 均为方阵，证明：尽管 AB 与 BA 可能不相等，但 $\det AB=\det BA$ 总是成立的.
34. 若 A，P 均为方阵，P 可逆，证明 $\det(PAP^{-1})=\det A$.
35. 假设 U 为方阵，满足 $U^{\mathrm{T}}U=I$，证明 $\det U=\pm 1$.
36. 假设 A 为方阵，且 $\det A^4=0$，解释为什么 A 不能是可逆的.

在习题37和习题38中，验证 $\det AB=(\det A)(\det B)$.（不要使用定理 6.）

37. $A=\begin{bmatrix} 3 & 0 \\ 6 & 1 \end{bmatrix}$, $B=\begin{bmatrix} 2 & 0 \\ 5 & 4 \end{bmatrix}$
38. $A=\begin{bmatrix} 3 & 6 \\ -1 & -2 \end{bmatrix}$, $B=\begin{bmatrix} 4 & 2 \\ -1 & -1 \end{bmatrix}$
39. 若 A,B 均为 3×3 矩阵，且 $\det A=4$，$\det B=-3$，利用（在课文中或上面习题中的）行列式性质计算
    a. $\det AB$　b. $\det 5A$　c. $\det B^{\mathrm{T}}$
    d. $\det A^{-1}$　e. $\det A^3$
40. 若 A,B 均为 4×4 矩阵，$\det A=-1$，$\det B=2$，计算
    a. $\det AB$　b. $\det B^5$　c. $\det 2A$
    d. $\det A^{\mathrm{T}}A$　e. $\det B^{-1}AB$
41. 验证 $\det A=\det B+\det C$，这里
    $A=\begin{bmatrix} a+e & b+f \\ c & d \end{bmatrix}$, $B=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $C=\begin{bmatrix} e & f \\ c & d \end{bmatrix}$
42. $A=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $B=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$，证明 $\det(A+B)=\det A+\det B$，当且仅当 $a+d=0$.
43. 验证 $\det A=\det B+\det C$，这里
    $$A=\begin{bmatrix} a_{11} & a_{12} & u_1+v_1 \\ a_{21} & a_{22} & u_2+v_2 \\ a_{31} & a_{32} & u_3+v_3 \end{bmatrix},$$
    $$B=\begin{bmatrix} a_{11} & a_{12} & u_1 \\ a_{21} & a_{22} & u_2 \\ a_{31} & a_{32} & u_3 \end{bmatrix}, C=\begin{bmatrix} a_{11} & a_{12} & v_1 \\ a_{21} & a_{22} & v_2 \\ a_{31} & a_{32} & v_3 \end{bmatrix}$$
    注意 A 与 $B+C$ 并不相同.
44. 矩阵 A 右乘一个初等矩阵 E 对 A 的列的影响与左乘一个初等矩阵对行的影响相同. 使用定理 5 和定理 3 以及 E^{T} 仍为初等矩阵这个明显的事实证明 $\det AE=(\det E)(\det A)$，要求不能使用定理 6.
45. [M]对几个随机选取的 4×5 矩阵和 5×6 矩阵，计算 $\det A^{\mathrm{T}}A$ 和 $\det AA^{\mathrm{T}}$. 当 A 的列数大于行数时，对 $A^{\mathrm{T}}A$ 和 AA^{T} 会是如何？
46. [M]如果 $\det A$ 接近于零，矩阵 A 是接近奇异的

吗？对 2.3 节习题 9 中接近奇异的 4×4 矩阵 A 进行实验，计算 A，$10A$，$0.1A$ 的行列式. 作为对比，计算这些矩阵的条件数. 当 A 为 4×4 单位矩阵时，重复这些计算，讨论你的结论.

练习题答案

1. 做行倍加变换，在第一列生成零，接着产生一个零行.

$$\begin{vmatrix} 1 & -3 & 1 & -2 \\ 2 & -5 & -1 & -2 \\ 0 & -4 & 5 & 1 \\ -3 & 10 & -6 & 8 \end{vmatrix} = \begin{vmatrix} 1 & -3 & 1 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & -4 & 5 & 1 \\ 0 & 1 & -3 & 2 \end{vmatrix} = \begin{vmatrix} 1 & -3 & 1 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & -4 & 5 & 1 \\ 0 & 0 & 0 & 0 \end{vmatrix} = 0$$

2. $$\det[\boldsymbol{v}_1 \ \ \boldsymbol{v}_2 \ \ \boldsymbol{v}_3] = \begin{vmatrix} 5 & -3 & 2 \\ -7 & 3 & -7 \\ 9 & -5 & 5 \end{vmatrix} = \begin{vmatrix} 5 & -3 & 2 \\ -2 & 0 & -5 \\ 9 & -5 & 5 \end{vmatrix} \quad \text{第 1 行加到第 2 行}$$

$$= -(-3)\begin{vmatrix} -2 & -5 \\ 9 & 5 \end{vmatrix} - (-5)\begin{vmatrix} 5 & 2 \\ -2 & -5 \end{vmatrix} \quad \text{按第 2 列的余因子展开}$$

$$= 3\cdot(35) + 5\cdot(-21) = 0$$

由定理 4，矩阵 $[\boldsymbol{v}_1 \ \ \boldsymbol{v}_2 \ \ \boldsymbol{v}_3]$ 不可逆，由可逆矩阵定理，这三列是线性相关的.

3.3 克拉默法则、体积和线性变换

本节应用前面几节的理论得到一些重要且有理论意义的公式以及行列式的几何解释.

克拉默法则

克拉默法则在各种理论计算中是必需的，例如，它被用来研究 $A\boldsymbol{x} = \boldsymbol{b}$ 的解受 $\boldsymbol{b}$ 中数值的变化而受到什么样的影响. 然而，这个公式对手算是没有多大效果的，除非是 2×2 或 3×3 矩阵.

对任意 $n \times n$ 矩阵 A 和任意的 $\mathbb{R}^n$ 中向量 $\boldsymbol{b}$，令 $A_i(\boldsymbol{b})$ 表示 A 中第 i 列由向量 $\boldsymbol{b}$ 替换得到的矩阵

$$A_i(\boldsymbol{b}) = [\boldsymbol{a}_1 \ \cdots \ \underset{\substack{\uparrow \\ \text{第 } i \text{ 列}}}{\boldsymbol{b}} \ \cdots \ \boldsymbol{a}_n]$$

定理 7 （克拉默法则）

设 A 是一个可逆的 $n \times n$ 矩阵，对 $\mathbb{R}^n$ 中任意向量 $\boldsymbol{b}$，方程 $A\boldsymbol{x} = \boldsymbol{b}$ 的唯一解可由下式给出

$$x_i = \frac{\det A_i(\boldsymbol{b})}{\det A}, i = 1, 2, \cdots, n \qquad (1)$$

证 用 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ 表示 A 的列，用 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 表示 $n \times n$ 单位阵 I 的列. 若 $A\boldsymbol{x} = \boldsymbol{b}$，由矩阵乘法的定义有

$$A \cdot I_i(\boldsymbol{x}) = A[\boldsymbol{e}_1 \ \cdots \ \boldsymbol{x} \ \cdots \boldsymbol{e}_n] = [A\boldsymbol{e}_1 \ \cdots \ A\boldsymbol{x} \ \cdots \ A\boldsymbol{e}_n]$$
$$= [\boldsymbol{a}_1 \cdots \ \boldsymbol{b} \ \cdots \boldsymbol{a}_n] = A_i(\boldsymbol{b})$$

由行列式的乘法性质

$$(\det A)(\det I_i(\boldsymbol{x})) = \det A_i(\boldsymbol{b})$$

左边第二个行列式为 x_i（沿第 i 行作余因子展开），从而 $(\det A)x_i=\det A_i(\boldsymbol{b})$. 由于 A 可逆，从而 $\det A\neq 0$，于是证明了(1)式. ■

例 1 利用克拉默法则解以下方程组

$$3x_1-2x_2=6$$
$$-5x_1+4x_2=8$$

解 视此方程组为 $A\boldsymbol{x}=\boldsymbol{b}$ 型，利用上面引入的记号，

$$A=\begin{bmatrix}3 & -2\\ -5 & 4\end{bmatrix},\ A_1(\boldsymbol{b})=\begin{bmatrix}6 & -2\\ 8 & 4\end{bmatrix},\ A_2(\boldsymbol{b})=\begin{bmatrix}3 & 6\\ -5 & 8\end{bmatrix}$$

由于 $\det A=2$，此方程组有唯一解. 由克拉默法则，有

$$x_1=\frac{\det A_1(\boldsymbol{b})}{\det A}=\frac{24+16}{2}=20$$
$$x_2=\frac{\det A_2(\boldsymbol{b})}{\det A}=\frac{24+30}{2}=27$$ ■

在工程上的应用

许多工程上的问题，特别是在电子工程和控制论，能用拉普拉斯变换进行分析，这种技巧将一个适当的线性微分方程组转变为一个线性代数方程组，它的系数含有一个参数.下一个例子说明可以出现的代数方程组类型.

例 2 考虑下列方程组，其中 s 是一个未定的参数，确定 s 的值，使得这个方程组有唯一解，利用克拉默法则写出这个解.

$$3sx_1-2x_2=4$$
$$-6x_1+sx_2=1$$

解 视此方程组为 $A\boldsymbol{x}=\boldsymbol{b}$ 型，则

$$A=\begin{bmatrix}3s & -2\\ -6 & s\end{bmatrix},\ A_1(\boldsymbol{b})=\begin{bmatrix}4 & -2\\ 1 & s\end{bmatrix},\ A_2(\boldsymbol{b})=\begin{bmatrix}3s & 4\\ -6 & 1\end{bmatrix}$$

由于 $\det A=3s^2-12=3(s+2)(s-2)$，当且仅当 $s\neq\pm 2$ 时，这个方程组有唯一解. 对这样的 s，方程组的解为 (x_1,x_2)，其中

$$x_1=\frac{\det A_1(\boldsymbol{b})}{\det A}=\frac{4s+2}{3(s+2)(s-2)}$$
$$x_2=\frac{\det A_2(\boldsymbol{b})}{\det A}=\frac{3s+24}{3(s+2)(s-2)}=\frac{s+8}{(s+2)(s-2)}$$ ■

一个求 A^{-1} 的公式

克拉默法则可以容易地导出一个求 $n\times n$ 矩阵 A 的逆的一般公式. A^{-1} 的第 j 列是一个向量 $\boldsymbol{x}$，满足

$$A\boldsymbol{x}=\boldsymbol{e}_j$$

此处 $\boldsymbol{e}_j$ 是单位矩阵的第 j 列，$\boldsymbol{x}$ 的第 i 个数值是 A^{-1} 中 $(i,\ j)$ 位置的数值，由克拉默法则

$$\{A^{-1}\text{中}(i\,,j)\text{元素}\} = x_i = \frac{\det A_i(\boldsymbol{e}_j)}{\det A} \tag{2}$$

回想起 A_{ji} 表示 A 的子矩阵，它由 A 去掉第 j 行和第 i 行得到. $A_i(\boldsymbol{e}_j)$ 按第 i 列的余因子展开式为

$$\det A_i(\boldsymbol{e}_j) = (-1)^{i+j}\det A_{ji} = C_{ji} \tag{3}$$

这里 C_{ji} 是 A 的一个余因子. 由（2），A^{-1} 的 $(i\,,\,j)$ 元素等于余因子 C_{ji} 除以 $\det A$.（注意：C_{ji} 的下标是 $(i\,,\,j)$ 的颠倒.）于是

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix} \tag{4}$$

公式（4）右边的余因子的矩阵称为 A 的伴随矩阵，记为 $\operatorname{adj} A$.（伴随这个术语在后面的线性变换课程中还有另一层意思.）下一个定理简单复述（4）式.

定理 8（逆矩阵公式）

设 A 是一个可逆的 $n\times n$ 矩阵，则 $A^{-1} = \dfrac{1}{\det A}\operatorname{adj} A$.

例 3 求矩阵 $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix}$ 的逆.

解 9 个余因子为

$$C_{11} = +\begin{vmatrix} -1 & 1 \\ 4 & -2 \end{vmatrix} = -2\,,\quad C_{12} = -\begin{vmatrix} 1 & 1 \\ 1 & -2 \end{vmatrix} = 3\,,\quad C_{13} = +\begin{vmatrix} 1 & -1 \\ 1 & 4 \end{vmatrix} = 5$$

$$C_{21} = -\begin{vmatrix} 1 & 3 \\ 4 & -2 \end{vmatrix} = 14\,,\quad C_{22} = +\begin{vmatrix} 2 & 3 \\ 1 & -2 \end{vmatrix} = -7\,,\quad C_{23} = -\begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} = -7$$

$$C_{31} = +\begin{vmatrix} 1 & 3 \\ -1 & 1 \end{vmatrix} = 4\,,\quad C_{32} = -\begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = 1\,,\quad C_{33} = +\begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix} = -3$$

伴随矩阵是余因子的矩阵的转置（例如，C_{12} 去到 $(2\,,\,1)$ 位置），从而

$$\operatorname{adj} A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}$$

我们可以直接计算 $\det A$，但下列计算对上面的结果提供了一个检验并计算出 $\det A$ 的方法.

$$(\operatorname{adj} A)\cdot A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}\begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix} = \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{bmatrix} = 14I$$

由于 $(\operatorname{adj} A)\cdot A = 14I$，由定理 8 得 $\det A = 14$，且

$$A^{-1}=\frac{1}{14}\begin{bmatrix}-2 & 14 & 4\\ 3 & -7 & 1\\ 5 & -7 & -3\end{bmatrix}=\begin{bmatrix}-1/7 & 1 & 2/7\\ 3/14 & -1/2 & 1/14\\ 5/14 & -1/2 & -3/14\end{bmatrix}$$ ■

数值计算的注解 定理 8 主要用在理论上的计算. A^{-1}的公式使我们不用实际计算出A^{-1}就可以推导出A^{-1}的性质. 如果确实需要求A^{-1}，除了特殊情形外，2.2 节中的算法给出了一个计算A^{-1}的很好的方法.

克拉默法则也是一个理论工具，它习惯于用来研究当$\boldsymbol{b}$或A中某数值改变时，$A\boldsymbol{x}=\boldsymbol{b}$的解如何敏感地变化.（可能由于当取得$\boldsymbol{b}$或$A$中的值时存在误差）. 当$A$是一个具有复数数值的 3×3 矩阵时，因为涉及复数的四则计算，$[A\ \ \boldsymbol{b}]$的行化简可能很麻烦，而行列式的计算相对容易，故克拉默法则有时在手算时被选用. 对一个很大的$n\times n$（实数或复数）矩阵，克拉默法则是无效的，仅计算一个行列式就大约与用行化简解$A\boldsymbol{x}=\boldsymbol{b}$有相同的工作量.

用行列式表示面积或体积

在下一个应用中，我们证明在本章介绍中阐述的行列式的几何解释. 尽管$\mathbb{R}^n$中的长度和距离的一般讨论直到第 6 章才给出，但我们在这里假设$\mathbb{R}^2$和$\mathbb{R}^3$中通常的欧几里得长度、面积、体积已经是清楚的.

定理 9 若A是一个 2×2 矩阵，则由A的列确定的平行四边形的面积为$|\det A|$，若A是一个 3×3 矩阵，则由A的列确定的平行六面体的体积为$|\det A|$.

证 若A为 2 阶对角矩阵，定理显然成立.

$$\left|\det\begin{bmatrix}a & 0\\ 0 & d\end{bmatrix}\right|=|ad|=\{\text{矩阵的面积}\}$$

见图 3-2. 若A不为对角情形，只需证$A=[\boldsymbol{a}_1\ \ \boldsymbol{a}_2]$能变换成一个对角矩阵，同时既不改变相应的平行四边形面积又不改变$|\det A|$. 由 3.2 节我们知道，当行列式的两列交换或一列的倍数加到另一列上时，行列式的绝对值不改变. 同时容易看到，这样的运算足以能够使A变换成对角矩阵. 由于列交换一点都不改变对应的平行四边形，所以只需证明下列在$\mathbb{R}^2$和$\mathbb{R}^3$中的向量的简单的几何现象就足够了.

图 3-2 面积$=|ad|$

> 设$\boldsymbol{a}_1$和$\boldsymbol{a}_2$为非零向量，则对任意数c，由$\boldsymbol{a}_1$和$\boldsymbol{a}_2$确定的平行四边形的面积等于由$\boldsymbol{a}_1$和$\boldsymbol{a}_2+c\boldsymbol{a}_1$确定的平行四边形的面积.

为了证明这个结论，我们可以假设$\boldsymbol{a}_2$不是$\boldsymbol{a}_1$的倍数，否则这两个平行四边形将退化成面积为 0. 若L是通过$\boldsymbol{0}$和$\boldsymbol{a}_1$的直线，则$\boldsymbol{a}_2+L$是通过$\boldsymbol{a}_2$且平行于L的直线，$\boldsymbol{a}_2+c\boldsymbol{a}_1$在此直线上，见图 3-3 点$\boldsymbol{a}_2$和$\boldsymbol{a}_2+c\boldsymbol{a}_1$到直线$L$具有相同的垂直距离，因此图 3-3 中的两个平行四边形具有相同的底边，即由$\boldsymbol{0}$到$\boldsymbol{a}_1$的线段，所以这两个平行四边形具有相同的面积，这就完成了$\mathbb{R}^2$的情形的证明.

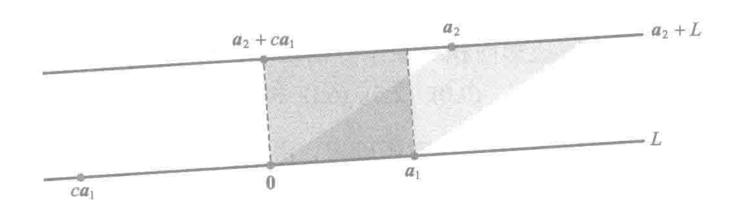


图 3-3　两个等面积的平行四边形

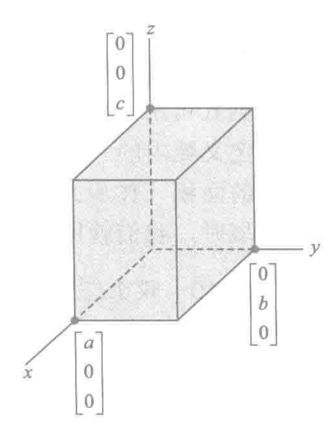


图 3-4　体积=$|a \ \ b \ \ c|$

类似可证明 $\mathbb{R}^3$ 的情形. 定理对 3×3 对角阵显然成立，见图 3-4. 任意一个 3×3 矩阵 A 均可用不改变$|\det A|$的列变换变换成对角矩阵（考虑在 A^{T} 上做行变换），所以只需证明这些变换不影响由 A 的列确定的平行六面体的体积就行了.

图 3-5 中用阴影给出一个平行六面体，具有两个倾斜的侧面，其体积等于在平面 $\mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_3\}$ 上的底面的面积乘以 $\boldsymbol{a}_2$ 到 $\mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_3\}$ 的垂直距离. 任意向量 $\boldsymbol{a}_2 + c\boldsymbol{a}_1$ 到 $\mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_3\}$ 的距离与 $\boldsymbol{a}_2$ 到 $\mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_3\}$ 的距离相同，这是因为 $\boldsymbol{a}_2 + c\boldsymbol{a}_1$ 位于平行于 $\mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_3\}$ 的平面 $\boldsymbol{a}_2 + \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_3\}$ 上，所以当$[\boldsymbol{a}_1 \ \ \boldsymbol{a}_2 \ \ \boldsymbol{a}_3]$变到$[\boldsymbol{a}_1 \ \ \boldsymbol{a}_2 + c\boldsymbol{a}_1 \ \ \boldsymbol{a}_3]$时，平行六面体的体积不变，于是列的倍加变换不影响平行六面体的体积. 由于列的交换也不影响体积，所以定理证毕.

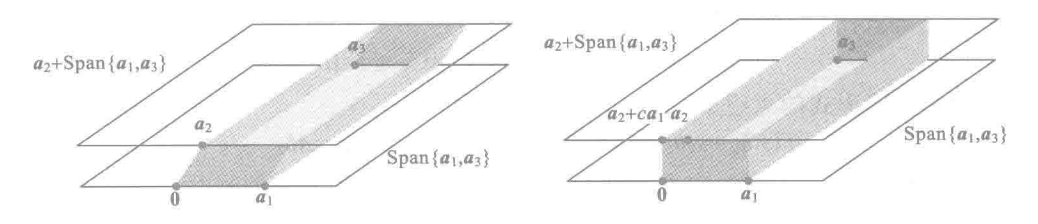


图 3-5　两个等体积的平行六面体

■

例 4　计算由点$(-2,2)$,$(0,3)$,$(4,-1)$和$(6,4)$确定的平行四边形的面积. 见图 3-6a.

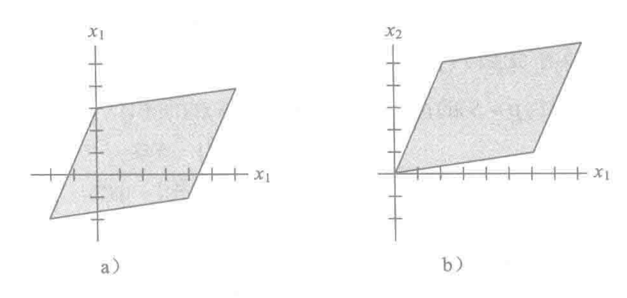


图 3-6　平移一个平行四边形不改变其面积

解 先将此平行四边形平移到使原点作为其一顶点的情形. 例如，将每个顶点坐标减去顶点 $(2,-2)$，这样，新的平行四边形面积与原平行四边形面积相同，其顶点为

$$(0,0),\ (2,5),\ (6,1)\ \text{和}\ (8,6)$$

见图 3-6b. 此平行四边形由 $A=\begin{bmatrix}2 & 6\\ 5 & 1\end{bmatrix}$ 的列所确定，由于 $|\det A|=|-28|$，所以所求的平行四边形的面积为 28. ■

线性变换

行列式可用于描述平面和 $\mathbb{R}^3$ 中线性变换的一个重要的几何性质. 若 T 是一个线性变换，S 是 T 的定义域内的一个集合，用 $T(S)$ 表示 S 中点的像集. 我们对 $T(S)$ 的面积（体积）与原来的集合 S 的面积（体积）相对比有何变化这件事感兴趣. 为了方便，当 S 是一个边界为平行四边形的区域时，我们就用 S 表示一个平行四边形.

定理 10 设 $T:\mathbb{R}^2\to\mathbb{R}^2$ 是由一个 2×2 矩阵 A 确定的线性变换，若 S 是 $\mathbb{R}^2$ 中一个平行四边形，则

$$\{T(S)\text{的面积}\}=|\det A|\cdot\{S\text{的面积}\} \tag{5}$$

若 T 是一个由 3×3 矩阵 A 确定的线性变换，而 S 是 R^3 中的一个平行六面体，则

$$\{T(S)\text{的体积}\}=|\det A|\cdot\{S\text{的体积}\} \tag{6}$$

证 考虑 2×2 的情形，$A=[\boldsymbol{a}_1\ \ \boldsymbol{a}_2]$. 一个顶点在 $\mathbb{R}^2$ 中原点的平行四边形，若它由向量 $\boldsymbol{b}_1$ 和 $\boldsymbol{b}_2$ 确定，则有以下等式：

$$S=\{s_1\boldsymbol{b}_1+s_2\boldsymbol{b}_2 : 0\leqslant s_1\leqslant 1,\ 0\leqslant s_2\leqslant 1\}$$

S 在 T 下的像由以下形式的点组成：

$$T(s_1\boldsymbol{b}_1+s_2\boldsymbol{b}_2)=s_1T(\boldsymbol{b}_1)+s_2T(\boldsymbol{b}_2)=s_1A\boldsymbol{b}_1+s_2A\boldsymbol{b}_2$$

这里 $0\leqslant s_1\leqslant 1, 0\leqslant s_2\leqslant 1$，从而 $T(S)$ 是一个由矩阵 $[A\boldsymbol{b}_1\ \ A\boldsymbol{b}_2]$ 的列确定的平行四边形. 这个矩阵可写成 AB，这里 $B=[\boldsymbol{b}_1\ \ \boldsymbol{b}_2]$. 由定理 9 和行列式的乘积定理，

$$\begin{aligned}\{T(S)\text{的面积}\}&=|\det AB|=|\det A|\cdot|\det B|\\&=|\det A|\cdot\{S\text{的面积}\}\end{aligned} \tag{7}$$

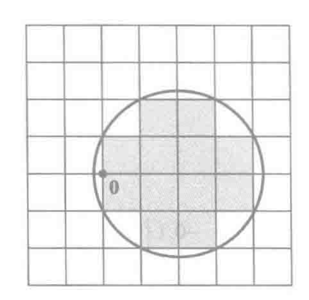
对任一个具有形式 $\boldsymbol{p}+S$ 的平行四边形，这里 $\boldsymbol{p}$ 是一个向量，S 是一个和上面一样有一个顶点在原点的平行四边形. 已知 T 将 $\boldsymbol{p}+S$ 变到 $T(\boldsymbol{p})+T(S)$（见习题 26.）由于平移不改变一个集合的面积，所以

$$\begin{aligned}\{T(\boldsymbol{p}+S)\text{的面积}\}&=\{T(\boldsymbol{p})+T(S)\text{的面积}\}\\&=\{T(S)\text{的面积}\}\quad\text{平移}\\&=|\det A|\cdot\{S\text{的面积}\}\quad\text{由(7)}\\&=|\det A|\cdot\{\boldsymbol{p}+S\text{的面积}\}\quad\text{平移}\end{aligned}$$

从而(5)式对 $\mathbb{R}^2$ 中任意的平行四边形均成立. (6)式在 3×3 情形的证明是类似的. ■

当我们尝试把定理 10 推广到 $\mathbb{R}^2$ 和 $\mathbb{R}^3$ 中不是由直线或平面所围的区域时，必须面对如何定

义和计算这个区域的面积或体积的问题. 这是一个在微积分中研究的问题，我们仅仅对 $\mathbb{R}^2$ 情形给出基本思想. 若 R 是一个有限面积的平面区域，则 R 可由其内部的小正方形方格来近似，通过使这些正方形充分小， R 的面积可由这些小正方形的面积之和充分地逼近. 见图 3-7.



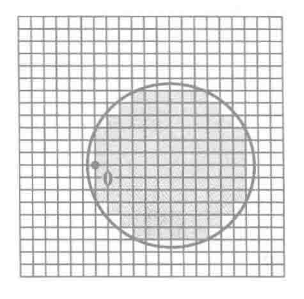


图 3-7 由正方形近似一个平面区域. 近似的程度随格子变细小而改进

若 T 是一个由 2×2 矩阵确定的线性变换，则 R 在 T 下的像由 R 中小正方形的像来近似. 定理 10 的证明说明每一个这样的像是一个平行四边形，它的面积是 $|\det A|$ 乘以这个正方形的面积. 若 R' 是 R 中这些正方形的并集，则 $T(R')$ 的面积是 $|\det A|$ 乘以 R' 的面积，见图 3-8. $T(R')$ 的面积也逼近 $T(R)$ 的面积. 当这个过程无限进行下去时，就可以得到下面推广的定理 10 的证明.

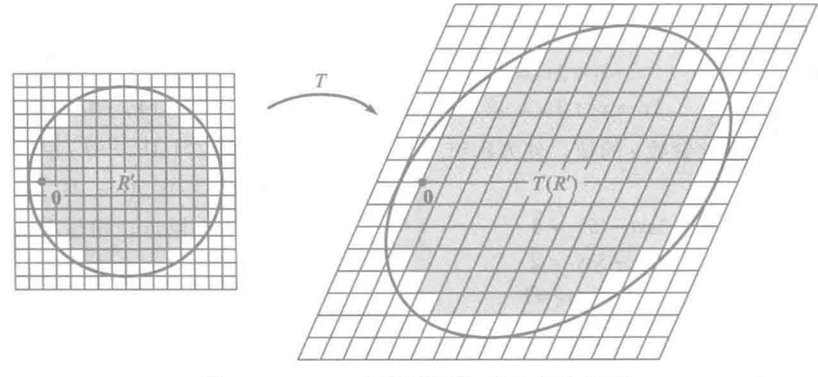
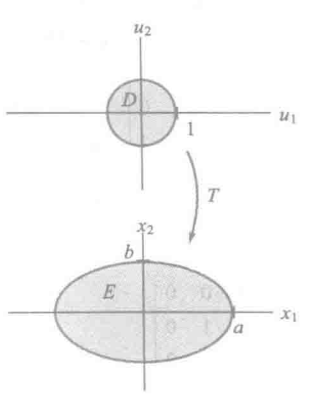


图 3-8 $T(R)$ 由平行四边形的并集近似



定理 10 的结论对 $\mathbb{R}^2$ 中任意具有有限面积的区域或 $\mathbb{R}^3$ 中具有有限体积的区域均成立.

例 5 若 a ,b 是正数，求由方程为 $\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}=1$ 的椭圆为边界的区域 E 的面积.

解 我们断言 E 是单位圆盘 D 在线性变换 T 下的像. 这里 T 由矩阵 $A=\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ 确定，这是因为若 $\boldsymbol{u}=\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ ，$\boldsymbol{x}=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ ，且 $\boldsymbol{x}=A\boldsymbol{u}$ ，则

$$u_1 = \frac{x_1}{a}, \quad u_2 = \frac{x_2}{b}$$

从而得 $\boldsymbol{u}$ 在此单位圆内，即满足 $u_1^2 + u_2^2 \leqslant 1$，当且仅当 $\boldsymbol{x}$ 在 E 内，即满足 $(x_1/a)^2 + (x_2/b)^2 \leqslant 1$. 由定理 10 的推广，

$$\begin{aligned}\{\text{椭圆的面积}\} &= \{T(D)\text{的面积}\} \\ &= |\det A| \cdot \{D\text{的面积}\} \\ &= a \cdot b \cdot \pi \cdot (1)^2 = \pi ab\end{aligned}$$

■

练习题

设 S 为由向量 $\boldsymbol{b}_1 = \begin{bmatrix}1\\3\end{bmatrix}$ 和 $\boldsymbol{b}_2 = \begin{bmatrix}5\\1\end{bmatrix}$ 确定的平行四边形，令 $A = \begin{bmatrix}1 & -0.1\\0 & 2\end{bmatrix}$. 计算 S 在映射 $\boldsymbol{x} \to A\boldsymbol{x}$ 下的像的面积.

习题 3.3

利用克拉默法则求习题 1~6 中方程组的解.

1. $\begin{aligned}5x_1 + 7x_2 &= 3\\ 2x_1 + 4x_2 &= 1\end{aligned}$
2. $\begin{aligned}4x_1 + x_2 &= 6\\ 5x_1 + 2x_2 &= 7\end{aligned}$
3. $\begin{aligned}3x_1 - 2x_2 &= 7\\ -5x_1 + 6x_2 &= -5\end{aligned}$
4. $\begin{aligned}-5x_1 + 3x_2 &= 9\\ 3x_1 - x_2 &= -5\end{aligned}$
5. $\begin{aligned}2x_1 + x_2 &= 7\\ -3x_1 + x_3 &= -8\\ x_2 + 2x_3 &= -3\end{aligned}$
6. $\begin{aligned}2x_1 + x_2 + x_3 &= 4\\ -x_1 + 2x_3 &= 2\\ 3x_1 + x_2 + 3x_3 &= -2\end{aligned}$

在习题 7~10 中，确定参数 s 的值，使方程组有唯一解，并求出该解.

7. $\begin{aligned}6sx_1 + 4x_2 &= 5\\ 9x_1 + 2sx_2 &= -2\end{aligned}$
8. $\begin{aligned}3sx_1 - 5x_2 &= 3\\ 9x_1 + 5sx_2 &= 2\end{aligned}$
9. $\begin{aligned}sx_1 - 2sx_2 &= -1\\ 3x_1 + 6sx_2 &= 4\end{aligned}$
10. $\begin{aligned}2sx_1 + x_2 &= 1\\ 3sx_1 + 6sx_2 &= 2\end{aligned}$

在习题 11~16 中，求给出矩阵的伴随矩阵，然后利用定理 8 给出矩阵的逆矩阵.

11. $\begin{bmatrix}0 & -2 & -1\\3 & 0 & 0\\-1 & 1 & 1\end{bmatrix}$
12. $\begin{bmatrix}1 & 1 & 3\\2 & -2 & 1\\0 & 1 & 0\end{bmatrix}$
13. $\begin{bmatrix}3 & 5 & 4\\1 & 0 & 1\\2 & 1 & 1\end{bmatrix}$
14. $\begin{bmatrix}3 & 6 & 7\\0 & 2 & 1\\2 & 3 & 4\end{bmatrix}$
15. $\begin{bmatrix}3 & 0 & 0\\-1 & 1 & 0\\-2 & 3 & 2\end{bmatrix}$
16. $\begin{bmatrix}1 & 2 & 4\\0 & -3 & 1\\0 & 0 & 3\end{bmatrix}$
17. 证明：若 A 是 2×2 矩阵，则对 A^{-1}，定理 8 与 2.2 节中定理 4 给出相同的公式.
18. 假设 A 中所有元素均为整数，且 $\det A = 1$. 解释为什么 A^{-1} 中所有元素也是整数.

在习题 19~22 中，分别求给定顶点的平行四边形的面积.

19. (0, 0), (5, 2), (6, 4), (11, 6)
20. (0, 0), (−1, 3), (4, −5), (3, −2)
21. (−1, 0), (0, 5), (1, −4), (2, 1)
22. (0, −2), (6, −1), (−3, 1), (3, 2)
23. 求一个顶点在原点，相邻顶点在 (1, 0, −2)，(1, 2, 4)，(7, 1, 0) 的平行六面体的体积.
24. 求一个顶点在原点，相邻顶点在 (1, 4, 0)，(−2, −5, 2)，(−1, 2, −1) 的平行六面体的体积.
25. 利用体积的概念解释为什么 3×3 矩阵 A 的行列式为零，当且仅当 A 不可逆. 要求不利用 3.2 节中的定理 4.（提示：考虑 A 的列.）
26. 令 $T:\mathbb{R}^m \to \mathbb{R}^n$ 为线性变换，$\boldsymbol{p}$ 为 $\mathbb{R}^m$ 中一个向量，S 为 $\mathbb{R}^m$ 中集合，证明：$\boldsymbol{p}+S$ 在 T 下的像等于 $\mathbb{R}^n$ 中平移的集合 $T(\boldsymbol{p})+T(S)$.
27. 令 S 是由向量 $\boldsymbol{b}_1 = \begin{bmatrix}-2\\3\end{bmatrix}$ 和 $\boldsymbol{b}_2 = \begin{bmatrix}-2\\5\end{bmatrix}$ 确定的平行四边形，$A = \begin{bmatrix}6 & -2\\-3 & 2\end{bmatrix}$，求 S 在映射 $\boldsymbol{x} \to A\boldsymbol{x}$

下像的面积.

28. 对 $\boldsymbol{b}_1=\begin{bmatrix}4\\-7\end{bmatrix}$，$\boldsymbol{b}_2=\begin{bmatrix}0\\1\end{bmatrix}$，$A=\begin{bmatrix}7&2\\1&1\end{bmatrix}$，重复 27 题.

29. 求 $\mathbb{R}^2$ 中顶点为 $\mathbf{0}$，$\boldsymbol{v}_1$，$\boldsymbol{v}_2$ 的三角形的面积公式.

30. 令 R 是顶点为 (x_1,y_1)，(x_2,y_2)，(x_3,y_3) 的三角形. 证明：$\{三角形的面积\}=\dfrac{1}{2}\left|\det\begin{bmatrix}x_1&y_1&1\\x_2&y_2&1\\x_3&y_3&1\end{bmatrix}\right|$.

（提示：通过减去一个顶点向量，将 R 平移到原点，再利用习题 29.）

31. 令 $T:\mathbb{R}^3\to\mathbb{R}^3$ 是由矩阵 $A=\begin{bmatrix}a&0&0\\0&b&0\\0&0&c\end{bmatrix}$ 确定的线性变换，这里 a，b，c 为正数，令 S 为单位球体，其边界曲面方程为 $x_1^2+x_2^2+x_3^2=1$.

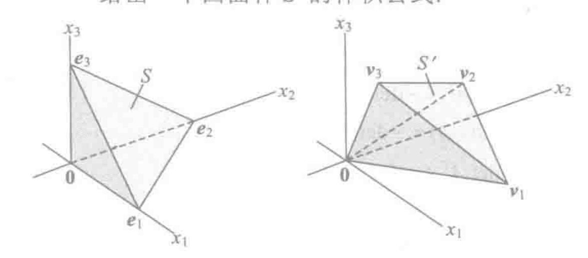
    a. 证明：$T(S)$ 的边界为椭球面，方程为 $\dfrac{x_1^2}{a^2}+\dfrac{x_2^2}{b^2}+\dfrac{x_3^2}{c^2}=1$.

    b. 利用单位球体体积为 $\dfrac{4}{3}\pi$ 的事实确定边界为(a)中椭球面的区域的体积.

32. 令 S 是 $\mathbb{R}^3$ 中具有顶点 $\mathbf{0}$，$\boldsymbol{e}_1$，$\boldsymbol{e}_2$，$\boldsymbol{e}_3$ 的四面体，S' 为具有顶点为 $\mathbf{0}$，$\boldsymbol{v}_1$，$\boldsymbol{v}_2$，$\boldsymbol{v}_3$ 的四面体，见下面的图形.

    a. 描述一个将 S 映射到 S' 上的线性变换.

    b. 利用 $\{S的体积\}=\dfrac{1}{3}\{底面积\}\cdot\{高\}$ 这个事实给出一个四面体 S' 的体积公式.



33. [M]任取一个 4×4 矩阵，检验定理 8 中的逆矩阵公式. 利用你的矩阵程序计算所有的 3×3 子矩阵的余因子，构造伴随矩阵，令 $B=(\operatorname{adj}A)/(\det A)$. 计算 $B-\operatorname{inv}(A)$，这里 $\operatorname{inv}(A)$ 是由矩阵程序计算出的 A 的逆. 使用浮点运算保留最大可能的小数位. 报告你的结论.

34. [M]任取一个 4×4 矩阵 A 和一个 4×1 向量 $\boldsymbol{b}$，检验克拉默法则. 求 $A\boldsymbol{x}=\boldsymbol{b}$ 的解并与 $A^{-1}\boldsymbol{b}$ 中元素相比较. 利用克拉默法则，对你的计算机程序，写出输出 $\boldsymbol{x}$ 中第二个元素的指令.

35. [M]在 MATLAB 中，对一个任意取的 30×30 的矩阵 A，使用 flops 命令统计出计算 A^{-1} 所需要的浮点运算的次数. 将这个数与计算 $(\operatorname{adj}A)/(\det A)$ 所需的浮点运算次数相比较.

练习题答案

S 的面积为 $\left|\det\begin{bmatrix}1&5\\3&1\end{bmatrix}\right|=14$，$\det A=2$，由定理 10，$S$ 在映射 $\boldsymbol{x}\to A\boldsymbol{x}$ 下的像的面积为

$$|\det A|\cdot\{S的面积\}=2\cdot 14=28.$$

补充习题

1. 指出每个命题的真假，给出理由.这里所有矩阵均设为方阵.

    a. 若 A 是 2×2 矩阵且其行列式为 0，则 A 的一列是另一列的倍数.

    b. 若 3×3 矩阵 A 的两行相等，则 $\det A=0$.

    c. 若 A 是 3×3 矩阵，则 $\det 5A=5\det A$.

    d. 若 A，B 均为 $n\times n$ 矩阵，满足 $\det A=2$，$\det B=3$，则 $\det(A+B)=5$.

    e. 若 A 为 $n\times n$ 矩阵且 $\det A=2$，则 $\det A^3=6$.

    f. 若 B 由 A 中交换两行生成，则 $\det B=\det A$.

g. 若 B 为由 A 中第 3 行乘以 5 生成，则 $\det B = 5\cdot\det A$.

h. 若 B 为通过 A 中任意 $n-1$ 行的线性组合加到另一行生成，则 $\det B=\det A$.

i. $\det A^{\mathrm{T}} = -\det A$.

j. $\det(-A) = -\det A$.

k. $\det A^{\mathrm{T}}A \geqslant 0$.

l. 任意一个 n 个未知数 n 个方程的方程组均可由克拉默法则解出.

m. 若 $\boldsymbol{u}, \boldsymbol{v}$ 属于 $\mathbb{R}^2$，且 $\det[\boldsymbol{u}\ \ \boldsymbol{v}]=10$，则平面中顶点为 $\boldsymbol{0}$，$\boldsymbol{u}, \boldsymbol{v}$ 的三角形面积为 10.

n. 若 $A^3=\boldsymbol{0}$，则 $\det A=0$.

o. 若 A 是可逆的，则 $\det A^{-1}=\det A$.

p. 若 A 是可逆的，则 $(\det A)(\det A^{-1})=1$.

使用行变换证明习题 2~4 中行列式都等于 0.

2. $\begin{vmatrix} 12 & 13 & 14 \\ 15 & 16 & 17 \\ 18 & 19 & 20 \end{vmatrix}$　　3. $\begin{vmatrix} 1 & a & b+c \\ 1 & b & a+c \\ 1 & c & a+b \end{vmatrix}$

4. $\begin{vmatrix} a & b & c \\ a+x & b+x & c+x \\ a+y & b+y & c+y \end{vmatrix}$

计算习题 5~6 中的行列式.

5. $\begin{vmatrix} 9 & 1 & 9 & 9 & 9 \\ 9 & 0 & 9 & 9 & 2 \\ 4 & 0 & 0 & 5 & 0 \\ 9 & 0 & 3 & 9 & 0 \\ 6 & 0 & 0 & 7 & 0 \end{vmatrix}$　　6. $\begin{vmatrix} 4 & 8 & 8 & 8 & 5 \\ 0 & 1 & 0 & 0 & 0 \\ 6 & 8 & 8 & 8 & 7 \\ 0 & 8 & 8 & 3 & 0 \\ 0 & 8 & 2 & 0 & 0 \end{vmatrix}$

7. 证明：$\mathbb{R}^2$ 中经过两不同点 (x_1, y_1), (x_2, y_2) 的直线的方程可以写成

$$\det\begin{bmatrix} 1 & x & y \\ 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \end{bmatrix}=0$$

8. 求一个类似于习题 7 的 3×3 行列式方程，描述通过点 (x_1, y_1) 且斜率为 m 的直线.

习题 9~10 涉及下列范德蒙德(Vandermonde)矩阵的行列式.

$$T=\begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix},\quad V(t)=\begin{bmatrix} 1 & t & t^2 & t^3 \\ 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \end{bmatrix}$$

9. 利用行变换证明 $\det T=(b-a)(c-a)(c-b)$.

10. 令 $f(t)=\det V$，x_1, x_2, x_3 两两不等，解释 $f(t)$ 为什么是一个三次多项式，证明 t^3 的系数非零，并在 f 的图上找三个点.

11. 求由点 $(1,4), (-1,5), (3,9)$ 和 $(5,8)$ 确定的平行四边形的面积. 你如何能判断由这些点确定的四边形刚好是一个平行四边形？

12. 利用平行四边形面积的概念描述对一个 2×2 矩阵 A 当且仅当 A 可逆时成立的命题.

13. 若 A 可逆，证明 $\operatorname{adj} A$ 也可逆，且 $(\operatorname{adj} A)^{-1}=\dfrac{1}{\det A}A$.

（提示：给定矩阵 B和C，什么计算能说明 C是B 的逆矩阵？）

14. 令 A, B, C, D 和 I 均为 $n\times n$ 矩阵，利用行列式的定义或性质证明下列公式，公式(c)在特征值（第 5 章）的应用中是有用的.

a. $\det\begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}=\det A$　b. $\det\begin{bmatrix} I & 0 \\ C & D \end{bmatrix}=\det D$

c. $\det\begin{bmatrix} A & 0 \\ C & D \end{bmatrix}=(\det A)(\det D)=\det\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$

15. 令 A, B, C, D 均为 $n\times n$ 矩阵，A 可逆.

a. 求矩阵 X和Y，使得生成分块 LU 的因式分解 $\begin{bmatrix} A & B \\ C & D \end{bmatrix}=\begin{bmatrix} I & 0 \\ X & I \end{bmatrix}\begin{bmatrix} A & B \\ 0 & Y \end{bmatrix}$.

然后证明

$$\det\begin{bmatrix} A & B \\ C & D \end{bmatrix}=(\det A)\cdot\det(D-CA^{-1}B)$$

b. 证明：若 $AC=CA$，则

$$\det\begin{bmatrix} A & B \\ C & D \end{bmatrix}=\det(AD-CB)$$

16. 设 J 是元素全为 1 的 $n\times n$ 矩阵，考虑 $A=(a-b)I+bJ$，即

$$A=\begin{bmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{bmatrix}$$

通过下面的计算推出 $\det A=(a-b)^{n-1}[a+(n-1)b]$：

a. 从第 2 行减去第 1 行，从第 3 行减去第 2 行，如此类推，说明为什么这样的操作不会改变矩阵的行列式.

b. 对（a）中所得的矩阵，将第 2 列加上第 1 列，将第 3 列加上第 2 列，如此类推，说明为什么这样的操作不会改变矩阵的行列式.

c. 求由（b）所得到的矩阵的行列式.

17\. 设 A 为 16 题中的矩阵，并设

$$B=\begin{bmatrix} a-b & b & b & \cdots & b \\ 0 & a & b & \cdots & b \\ 0 & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & b & b & \cdots & a \end{bmatrix},$$

$$C=\begin{bmatrix} b & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{bmatrix}$$

这里 A,B,C 基本相同，除了 A 的第一列等于 B 和 C 的第一列之和. 在 3.2 节讨论的行列式函数的线性性质给出 $\det A=\det B+\det C$. 利用这个事实用归纳法证明 16 题给出的公式.

18\. [M]运用 16 题的结论求下面矩阵的行列式，使用矩阵程序验证你的结果.

$$\begin{bmatrix} 3 & 8 & 8 & 8 \\ 8 & 3 & 8 & 8 \\ 8 & 8 & 3 & 8 \\ 8 & 8 & 8 & 3 \end{bmatrix} \quad \begin{bmatrix} 8 & 3 & 3 & 3 & 3 \\ 3 & 8 & 3 & 3 & 3 \\ 3 & 3 & 8 & 3 & 3 \\ 3 & 3 & 3 & 8 & 3 \\ 3 & 3 & 3 & 3 & 8 \end{bmatrix}$$

19\. [M]使用矩阵程序求下面矩阵的行列式

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}$$

使用上面计算结果猜测下面矩阵的行列式，用行变换计算其行列式来验证你的猜测.

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & n \end{bmatrix}$$

20\. [M]使用 19 题的方法猜测下面矩阵的行列式，验证你的猜测.（提示：利用习题 14(c)和习题 19 的结论.）

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 3 & 3 & \cdots & 3 \\ 1 & 3 & 6 & \cdots & 6 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 3 & 6 & \cdots & 3(n-1) \end{bmatrix}$$

第4章 向量空间

介绍性实例 空间飞行与控制系统

1981 年 4 月，在清凉的棕榈星期日（复活节前的星期日）的早晨，12 层楼高，75 吨重的美国哥伦比亚号航天飞机壮观地飞离发射台. 作为十年集中研究和开发的成果，这架美国第一架航天飞机是控制系统工程设计的最好的例子，涉及工程学的许多分支——航空、化学、电子、液压以及机械工程.

航天飞机的控制系统对飞行是绝对关键的. 由于航天飞机是一个不稳定的空中机体，在大气层飞行时它需要不间断地用计算机监控. 飞行控制系统不断地向空气动力控制表面和 44 个小推进器喷口发送命令. 图 4-1 展示了一个典型的闭环反馈系统，它用来控制航天飞机在飞行时的俯仰角（即头部圆锥体的仰角）. 符号 ⊗ 表示这里来自各种各样的传感器的信号被添加到沿此图顶部传递的计算机信号中.

从数学的角度看，一个工程学系统输入和输出信号都是函数，如图 4-1 所示，这些函数的加法和标量乘法在应用中是重要的. 在 4.1 节和 4.8 节中，我们将会看到，函数的这两个运算具有完全类似于 $\mathbb{R}^n$ 中向量的加法和标量乘法的代数性质. 由于这个原因，所有可能的输入（函数）的集合称为一个向量空间. 系统工程学的数学基础依赖于函数的向量空间，因此我们需要把 $\mathbb{R}^n$ 中向量的理论推广到包括这些函数. 后面，我们将看到其他向量空间如何出现在工程学、物理学和统计学中.

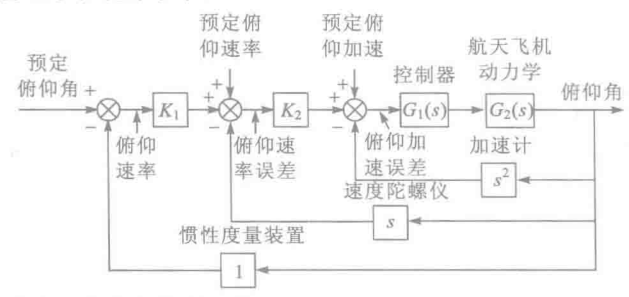


图 4-1 航天飞机的俯仰角控制系统（来源：*Control Systems Engineering*, by Norman S.Nise, Benjamin-Cummings Publishing, 1992, p.274. Simplified drawing based on *Space Shuttle GN&C Operations Manual,* Rockwell International,1988.）

在第 1 章和第 2 章中种植的数学种子在这一章中生长并且开始开花. 当我们把 $\mathbb{R}^n$ 仅仅看作自然地出现在应用问题中的各种向量空间之一时，线性代数的美和力量将更清楚地显露出来. 我们将看到从 $\mathbb{R}^n$ 本身研究向量空间并不是很困难的，因为我们可以利用 $\mathbb{R}^2$ 和 $\mathbb{R}^3$ 中的几何经验使许多几何概念直观化.

在 4.1 节中给出一些基本定义之后，一般向量空间的框架逐步展开并贯穿于全章. 4.3~4.5 节的一个目标是表明其他的向量空间是如何类似于 $\mathbb{R}^n$ 的. 4.6 节讨论的秩是本章的高潮之一，利用向量空间的术语将矩阵的重要知识连在一起. 4.8 节将本章的理论应用到离散信号和差分方程，它们在比在航天飞机中用到的数字控制系统中有用. 4.9 节中的马尔可夫（Markov）链与本章中理论性较强的部分相比有了质的变化，它为第 5 章中出现的概念提供了较好的例子.

4.1 向量空间与子空间

第 1 章和第 2 章中许多内容停留在 $\mathbb{R}^n$ 的一些简单且明显的性质上，这些性质列在 1.3 节中. 同样，我们将看到许多其他的数学系统具有与 $\mathbb{R}^n$ 中相同的性质，这些具体而有趣的性质列在下面定义中.

定义 一个向量空间是由一些被称为向量的对象构成的非空集合 V，在这个集合上定义两个运算，称为加法和标量乘法（标量取实数），服从以下公理（或法则）[⊖]，这些公理必须对 V 中所有向量 $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}$ 及所有标量 c 和 d 均成立.

1. $\boldsymbol{u},\boldsymbol{v}$ 之和表示为 $\boldsymbol{u}+\boldsymbol{v}$，仍在 V 中.
2. $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$.
3. $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$.
4. V 中存在一个零向量 $\boldsymbol{0}$，使得 $\boldsymbol{u}+\boldsymbol{0}=\boldsymbol{u}$.
5. 对 V 中每个向量 $\boldsymbol{u}$，存在 V 中向量 $-\boldsymbol{u}$，使得 $\boldsymbol{u}+(-\boldsymbol{u})=\boldsymbol{0}$.
6. $\boldsymbol{u}$ 与标量 c 的标量乘法记为 $c\boldsymbol{u}$，仍在 V 中.
7. $c(\boldsymbol{u}+\boldsymbol{v})=c\boldsymbol{u}+c\boldsymbol{v}$.
8. $(c+d)\boldsymbol{u}=c\boldsymbol{u}+d\boldsymbol{u}$.
9. $c(d\boldsymbol{u})=(cd)\boldsymbol{u}$.
10. $1\boldsymbol{u}=\boldsymbol{u}$.

只需利用这些公理，我们可以证明公理 4 中的零向量是唯一的. 对 V 中每个向量 $\boldsymbol{u}$，公理 5 中向量 $-\boldsymbol{u}$ 称为 $\boldsymbol{u}$ 的**负向量**，也是唯一的，见习题 25 和 26，下列简单事实的证明也在习题中给出了大致轮廓：

⊖ 更专业地讲，V 是一个实数向量空间，本章所有理论对复向量空间（数取复数）也成立，在第 5 章中我们会看到一些例子. 到现在为止，所有数假设为实数.

对 V 中每个向量 $\boldsymbol{u}$ 和任意标量 c，有

$$0\boldsymbol{u}=\boldsymbol{0} \tag{1}$$

$$c\boldsymbol{0}=\boldsymbol{0} \tag{2}$$

$$-\boldsymbol{u}=(-1)\boldsymbol{u} \tag{3}$$

例 1 空间 $\mathbb{R}^n$（$n\geqslant 1$）为向量空间的首要的例子. $\mathbb{R}^3$ 的几何直觉可以帮助我们使本章的许多概念清晰化并且直观化. ■

例 2 设 V 是三维空间中所有有向线段的集合，如果其中两个向量方向相同且长度相等，则二者视为相等，由平行四边形法则定义加法（见 1.3 节），且对 V 中每个向量 $\boldsymbol{v}$，定义 $c\boldsymbol{v}$ 为长度等于 $|c|$ 乘以 $\boldsymbol{v}$ 的长度. 若 $c\geqslant 0$，则 $c\boldsymbol{v}$ 与 $\boldsymbol{v}$ 同向，否则与 $\boldsymbol{v}$ 反向（见图 4-2），证明 V 是一个向量空间. 这个空间是物理学中变力问题中一个常用的模型.

解 V 的定义是按几何方式的定义的，只用到长度和方向的概念，没有涉及 xyz 坐标系，一个零长度的有向线段就是一个单点，表示零向量. $\boldsymbol{v}$ 的负向量为 $(-1)\boldsymbol{v}$，所以公理 1，4，5，6，10 显然成立，剩下的可用几何方法证明，比如图 4-3、图 4-4.

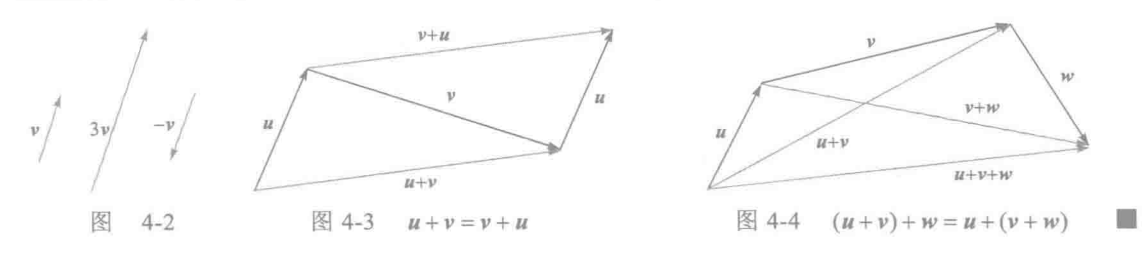


图 4-2　　图 4-3 $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$　　图 4-4 $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$ ■

例 3 设 $\mathbb{S}$ 是数的双向无穷序列空间（通常写成行而不写成列）：

$$\{y_k\}=(\cdots,y_{-2},y_{-1},y_0,y_1,y_2,\cdots)$$

若 $\{z_k\}$ 是 $\mathbb{S}$ 中的另一个元素，则和 $\{y_k\}+\{z_k\}$ 是序列 $\{y_k+z_k\}$，它由 $\{y_k\}$ 与 $\{z_k\}$ 对应项之和构成，数乘 $c\{y_k\}$ 是序列 $\{cy_k\}$，用与 $\mathbb{R}^n$ 中相同的方法可以证明向量空间的那些公理.

$\mathbb{S}$ 中的元素来源于工程学，例如，每当一个信号在离散时间上被测量（或被简化）时，它就可被看作 $\mathbb{S}$ 中的一个个元素，这样的信号可以是电的、机械的、光的等. 在本章“介绍性实例”中提到的航天飞机的大控制系统，就使用离散或“数字”信号. 为了方便，我们将称 $\mathbb{S}$ 为（离散时间）信号空间，一个信号可以直观地由图 4-5 表示.

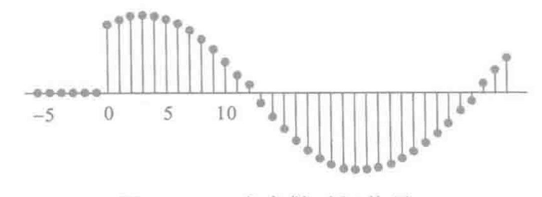


图 4-5 一个离散时间信号 ■

例 4 对 $n\geqslant 0$，次数最高为 n 的多项式集合 $\mathbb{P}_n$ 由形如下列的多项式组成：

$$\boldsymbol{p}(t)=a_0+a_1t+a_2t^2+\cdots+a_nt^n \tag{4}$$

这里系数 $a_0,\cdots,a_n$ 和变量 t 均为实数，p 的次数是（4）中系数不为零的项中 t 的最高幂，若 $\boldsymbol{p}(t)=a_0\neq 0$，则 p 的次数为零，若所有系数均为零，则 $\boldsymbol{p}$ 称为零多项式，零多项式包含在 $\mathbb{P}_n$ 中，尽管它的次数没有定义.

若 $\boldsymbol{p}$ 由（4）式给出，$\boldsymbol{q}(t)=b_0+b_1t+\cdots+b_nt^n$，则 $\boldsymbol{p}+\boldsymbol{q}$ 定义为：

$$\begin{aligned}(\boldsymbol{p}+\boldsymbol{q})(t)&=\boldsymbol{p}(t)+\boldsymbol{q}(t)\\&=(a_0+b_0)+(a_1+b_1)t+\cdots+(a_n+b_n)t^n\end{aligned}$$

标量乘法 $c\boldsymbol{p}$ 定义为：

$$(c\boldsymbol{p})(t)=c\boldsymbol{p}(t)=ca_0+(ca_1)t+\cdots+(ca_n)t^n$$

这些定义满足公理 1 和公理 6，这是因为 $\boldsymbol{p}+\boldsymbol{q}$ 和 $c\boldsymbol{p}$ 均为次数不超过 n 的多项式，公理 2，3 和公理 7~10 由实数性质验证. 显然，零多项式可以作为公理 4 中的零向量，最后，$(-1)\boldsymbol{p}$ 作为 $\boldsymbol{p}$ 的负向量，所以满足公理 5，于是 $\mathbb{P}_n$ 是一个向量空间.

对不同的 n，向量空间 $\mathbb{P}_n$ 用于，比如说，统计数据的趋势分析中，在 6.8 节中讨论. ■

例 5 设 V 是定义在集合 $\mathbb{D}$ 上的全体实函数的集合（典型地，$\mathbb{D}$ 为实数集或实轴上的区间）. 用通常方式定义加法（即 $\boldsymbol{f}+\boldsymbol{g}$）仍为函数，在 $\mathbb{D}$ 中 t 点的值为 $\boldsymbol{f}(t)+\boldsymbol{g}(t)$. 同样，对标量 c 和 V 中的 $\boldsymbol{f}$，标量乘法仍为函数，在 t 的值为 $c\boldsymbol{f}(t)$. 例如，若 $\mathbb{D}=\mathbb{R}$，$\boldsymbol{f}(t)=1+\sin 2t$，$\boldsymbol{g}(t)=2+0.5t$，则

$$(\boldsymbol{f}+\boldsymbol{g})(t)=3+\sin 2t+0.5t\ ,\ (2\boldsymbol{g})(t)=4+t$$

V 中两个函数相等，当且仅当对任意 $\mathbb{D}$ 中的 t 函数值相等. 从而 V 中的零向量是恒等于零的函数 $\boldsymbol{f}(t)=0$，任意 $t\in D$，f 的负向量为 $(-1)\boldsymbol{f}$. 公理 1 和公理 6 显然成立，其余公理由实数性质得证，所以 V 为一个向量空间. ■

把例 5 中的向量空间 V 中每个函数看作一个独立的个体是重要的，如同在向量空间中的一个“点”或向量那样，两个向量 $\boldsymbol{f}$ 与 $\boldsymbol{g}$ 的和（$\boldsymbol{f}$ 与 $\boldsymbol{g}$ 为 V 中的函数，即向量空间中的元素）可以通过图 4-6 给予直观解释，因为这样可以帮助你将 $\mathbb{R}^n$ 中建立的几何直觉上升到一般向量空间，“学习指南”可以帮你接受更一般的观点.

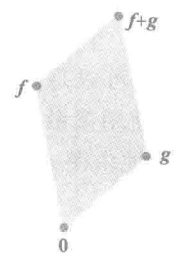


图 4-6 两个向量（函数）之和

子空间

在许多问题中，一个向量空间是由一个大的向量空间中适当的向量的子集所构成. 在此情形下，向量空间的 10 个公理中只需要验证三个，其余的自然成立.

定义 向量空间 V 的一个子空间是 V 的一个满足以下三个性质的子集 H：

a. V 中的零向量在 H 中.[㊀]

b. H 对向量加法封闭，即对 H 中任意向量 $\boldsymbol{u},\boldsymbol{v}$，和 $\boldsymbol{u}+\boldsymbol{v}$ 仍在 H 中.

c. H 对标量乘法封闭，即对 H 中任意向量 $\boldsymbol{u}$ 和任意标量 c，向量 $c\boldsymbol{u}$ 仍在 H 中.

性质(a)、(b)、(c)保证 V 的子空间 H 本身对 V 中的向量空间运算而言是一个向量空间. 为证此结论，注意到(a)、(b)、(c)分别是公理 1、4、6. 公理 2 和 3 及公理 7~10 自然成立，由于它们对 V 中所有元素均成立，自然包括 H 中的元素，公理 5 在 H 中也成立，因为若 $\boldsymbol{u}$ 在 H 中，则由(c)知，$(-1)\boldsymbol{u}$ 也在 H 中，同时，从本节等式（3）知，$(-1)\boldsymbol{u}$ 即公理 5 中的 $-\boldsymbol{u}$.

这样每个子空间都是一个向量空间. 反之，每个向量空间是一个子空间（针对本身或其他更大的空间而言）. 两个向量空间，若其中一个在另一个内部，此时子空间这个词被使用，而 V 的子空间是将 V 看作更大的空间（见图 4-7）.

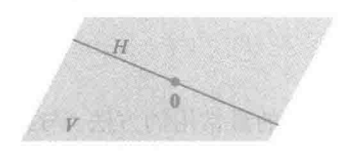


图 4-7　V 的一个子空间

例 6　向量空间 V 中仅由零向量组成的集合是 V 的一个子空间，称为零子空间，写成 $\{\mathbf{0}\}$. ■

例 7　令 $\mathbb{P}$ 为全体实系数多项式的集合，$\mathbb{P}$ 中运算的定义与函数运算相同，则 $\mathbb{P}$ 是定义在 $\mathbb{R}$ 上的全体实函数的一个子空间. 再者，对每个 $n\geqslant 0$, $\mathbb{P}_n$ 是 $\mathbb{P}$ 的子空间，因为 $\mathbb{P}_n$ 是 $\mathbb{P}$ 的子集，它包含零多项式，且 $\mathbb{P}_n$ 中两个多项式之和仍在 $\mathbb{P}_n$ 中，数乘以 $\mathbb{P}_n$ 中一个多项式仍在 $\mathbb{P}_n$ 中. ■

例 8　向量空间 $\mathbb{R}^2$ 不是 $\mathbb{R}^3$ 的子空间，因为 $\mathbb{R}^2$ 甚至不是 $\mathbb{R}^3$ 的子集（$\mathbb{R}^3$ 中的向量有 3 个分量，而 $\mathbb{R}^2$ 中的向量仅有两个分量），集合

$$H=\left\{\begin{bmatrix}s\\t\\0\end{bmatrix}: s,t\text{是实数}\right\}$$

是 $\mathbb{R}^3$ 的一个子集，尽管从逻辑上讲它与 $\mathbb{R}^2$ 不同，但看起来很像 $\mathbb{R}^2$，见图 4-8，证明 H 是 $\mathbb{R}^3$ 的一个子空间.

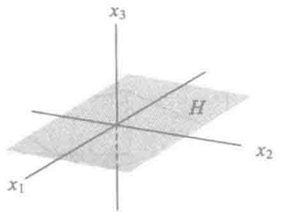


图 4-8　作为 $\mathbb{R}^3$ 的子空间的 x_1x_2 平面

㊀ 有些教材将定义中的(a)替换成 H 非空，则(a)可由(c)推出，且有 $0\boldsymbol{u}=\mathbf{0}$，但检验子空间最好的方法是首先观察零向量，若零向量在 H 中，则(b)、(c)必须验证，若零向量不在 H 中，则 H 不是子空间，且其余性质不必检验.

解　零向量在 H 中，且对向量的加法和标量乘法，H 是封闭的，这是因为对 H 中的向量而言，那些运算产生的向量中第 3 个分量仍为零（从而属于 H），所以 H 是 $\mathbb{R}^3$ 的一个子空间. ■

例 9　$\mathbb{R}^3$ 中一个不通过原点的平面不是 $\mathbb{R}^3$ 中子空间. 因为此平面不包含 $\mathbb{R}^3$ 中的零向量. 类似地，$\mathbb{R}^2$ 中一条不通过原点的直线，如图 4-9 所示，也不是 $\mathbb{R}^2$ 的子空间.

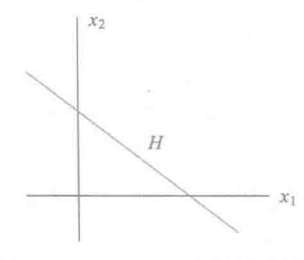


图 4-9　不是向量空间的直线 ■

由一个集合生成的子空间

下一个例子说明了一个描述子空间的最常用的方法. 与第 1 章中相同，**线性组合**这个词表示一些向量的任意标量乘法之和，$\text{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 表示所有可以表示成 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 的线性组合的向量集合.

例 10　给定向量空间 V 中向量 $\boldsymbol{v}_1,\boldsymbol{v}_2$，令 $H=\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$，证明 H 是 V 的一个子空间.

解　由于 $\mathbf{0}=0\boldsymbol{v}_1+0\boldsymbol{v}_2$，所以零向量在 H 中，为证 H 对加法封闭，任取 H 中两个向量，即

$$\boldsymbol{u}=s_1\boldsymbol{v}_1+s_2\boldsymbol{v}_2\ ,\ \boldsymbol{w}=t_1\boldsymbol{v}_1+t_2\boldsymbol{v}_2$$

对向量空间 V，由公理 2、3 和 8 知，

$$\begin{aligned}\boldsymbol{u}+\boldsymbol{w}&=(s_1\boldsymbol{v}_1+s_2\boldsymbol{v}_2)+(t_1\boldsymbol{v}_1+t_2\boldsymbol{v}_2)\\&=(s_1+t_1)\boldsymbol{v}_1+(s_2+t_2)\boldsymbol{v}_2\end{aligned}$$

所以 $\boldsymbol{u}+\boldsymbol{w}$ 在 H 中，进一步，若 c 是任意标量，由公理 7 和 9 知，

$$c\boldsymbol{u}=c(s_1\boldsymbol{v}_1+s_2\boldsymbol{v}_2)=(cs_1)\boldsymbol{v}_1+(cs_2)\boldsymbol{v}_2$$

这证明 $c\boldsymbol{u}$ 在 H 中，从而 H 对标量乘法封闭，从而 H 是 V 的一个子空间. ■

在 4.5 节中，我们将证明 $\mathbb{R}^3$ 的每一个非零子空间，除了 $\mathbb{R}^3$ 本身之外，要么是 $\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$，这里 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 是某线性无关的两个向量，要么是 $\text{Span}\{\boldsymbol{v}\}$，$\boldsymbol{v}\neq\mathbf{0}$. 对第一种情形，此子空间是一个通过原点的平面，第二种情形，子空间是一条通过原点的直线（见图 4-10）. 记住这些几何图形是有好处的，甚至对一个抽象的向量空间也有帮助.

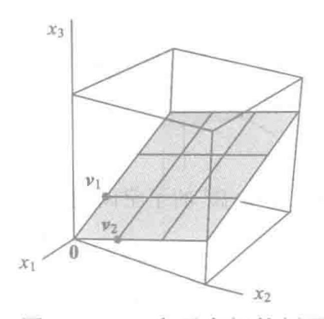


图 4-10　一个子空间的例子

例 10 的讨论可以很容易地推广从而证明下面的定理.

定理 1 若 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 在向量空间 V 中，则 $\text{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 是 V 的一个子空间.

我们称 $\text{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 是由 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ **生成（或张成）的子空间**，任给 V 的子空间 H，H 的**生成（或张成）集**是集合 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}\subset H$，使得 $H=\text{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$.

下例说明如何使用定理 1.

例 11 令 H 是所有形如 $(a-3b,b-a,a,b)$ 的向量的集合，这里 a,b 是任意数，即

$$H=\{(a-3b,b-a,a,b):a,b\in\mathbb{R}\}$$

证明 H 是 $\mathbb{R}^4$ 的一个子空间.

解 将 H 中向量写成列向量，则 H 中任意向量具有如下形式：

$$\begin{bmatrix} a-3b \\ b-a \\ a \\ b \end{bmatrix} = a\underset{\substack{\uparrow \\ \boldsymbol{v}_1}}{\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}} + b\underset{\substack{\uparrow \\ \boldsymbol{v}_2}}{\begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \end{bmatrix}}$$

这个计算表明 $H=\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$，这里 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 标示如上，从而由定理 1 知 H 是 $\mathbb{R}^4$ 的一个子空间. ■

例11 表明一个有用的技巧，用来表示作为某些向量线性组合的集合的子空间H. 若 $H=\text{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$，我们可以把这个生成集中的向量 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 看作“柄”，它们使我们能够掌握这个子空间 H. H 中无穷多个向量的运算经常被简化成生成集中的有限多个向量的运算.

例 12 问 h 取何值时，$\boldsymbol{y}$ 在由 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 生成的 $\mathbb{R}^3$ 的子空间中？其中

$$\boldsymbol{v}_1=\begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix},\ \boldsymbol{v}_2=\begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix},\ \boldsymbol{v}_3=\begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix},\ \boldsymbol{y}=\begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$$

解 此问题是 1.3 节中的练习题 2，这里用子空间这个词还不如用 $\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$. 此处 $\boldsymbol{y}$ 在 $\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 中当且仅当 $h=5$. 这个解法现在值得复习一下，可以复习 1.3 节中习题 11~16 和习题 19~21. ■

虽然本章中许多向量空间是 $\mathbb{R}^n$ 的子空间，但牢记那些抽象理论用在其他向量空间仍成立是很重要的，函数的向量空间起源于许多应用，这些空间在后面的章节将受到关注.

练习题

1. 通过证明标量乘法不封闭，证明：$\mathbb{R}^2$ 中所有形如 $(3s,2+5s)$ 的点集 H 不能构成一个向量空间.（H 中一个特殊向量 $\boldsymbol{u}$ 和数 c，使 $c\boldsymbol{u}$ 不在 H 中.）
2. 令 $W=\text{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$，这里 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 在向量空中 V 中，证明 $\boldsymbol{v}_k$ 在 W 中，$1\leqslant k\leqslant p$.（提示：先写一个能证明 $\boldsymbol{v}_1$ 在 V 中的方程，再类似证明一般情形.）

习题 4.1

1. 令 V 是 xy 平面中的第一象限，即
$$V=\left\{\begin{bmatrix}x\\y\end{bmatrix}:x\geqslant 0,y\geqslant 0\right\}$$
   a. 若 $\boldsymbol{u},\boldsymbol{v}$ 在 V 中，$\boldsymbol{u}+\boldsymbol{v}$ 在 V 中吗？为什么？
   b. 找 V 中一个特殊向量 $\boldsymbol{u}$ 和一个特殊数 c，使得 $c\boldsymbol{u}$ 不在 V 中.（这足以说明 V 不是一个向量空间.）
2. 令 W 是 xy 平面中第一象限与第三象限的并集，即 $W=\left\{\begin{bmatrix}x\\y\end{bmatrix}:xy\geqslant 0\right\}$.
   a. 若 $\boldsymbol{u}$ 在 W 中，c 为任意数，$c\boldsymbol{u}$ 在 W 中吗？为什么？
   b. 找 W 中特殊向量 $\boldsymbol{u},\boldsymbol{v}$，使得 $\boldsymbol{u}+\boldsymbol{v}$ 不在 W 中. 这足以说明 W 不是一个向量空间.
3. 令 H 是 xy 平面上所有单位圆内和圆上的点集，即 $H=\left\{\begin{bmatrix}x\\y\end{bmatrix}:x^2+y^2\leqslant 1\right\}$，找出一个特殊的例子，两个向量或者是一个向量与一个数，证明 H 不是 $\mathbb{R}^2$ 中的一个子空间.
4. 构造一个几何图形说明为什么 $\mathbb{R}^2$ 中一条不过原点的直线对向量的加法不封闭.

在习题 5~8 中，n 取适当的值，判定给出的集合是否为 $\mathbb{P}_n$ 的子空间，验证你的答案.

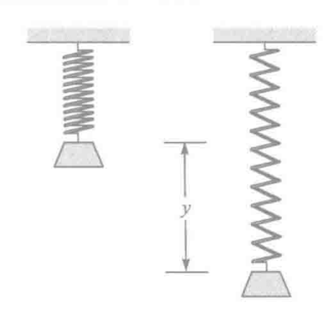
5. 所有形如 $p(t)=at^2$ 的多项式，此处 $a\in\mathbb{R}$.
6. 所有形如 $p(t)=a+t^2$ 的多项式，此处 $a\in\mathbb{R}$.
7. 所有次数最多是 3 的整系数多项式.
8. 所有 $\mathbb{P}_n$ 中使得 $p(0)=0$ 的多项式.
9. 令 H 为所有形如 $\begin{bmatrix}-2t\\5t\\3t\end{bmatrix}$ 的向量集，找 $\mathbb{R}^3$ 中一个向量 $\boldsymbol{v}$，使得 $H=\mathrm{Span}\{\boldsymbol{v}\}$，为什么这样能证明 H 是 $\mathbb{R}^3$ 的一个子空间？
10. 令 H 为所有形如 $\begin{bmatrix}2t\\0\\-7t\end{bmatrix}$ 的向量集，其中 t 是任意实数. 证明 H 是 $\mathbb{R}^3$ 的一个子空间.（利用习题 9 的方法.）
11. 令 W 是所有形如 $\begin{bmatrix}2b+3c\\-b\\2c\end{bmatrix}$ 的向量集，这里 b,c 是任意数，找向量 $\boldsymbol{u},\boldsymbol{v}$ 使得 $W=\mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$，为什么这样能证明 W 是 $\mathbb{R}^3$ 的一个子空间？
12. 令 W 是所有形如 $\begin{bmatrix}2s+4t\\2s\\2s-3t\\5t\end{bmatrix}$ 的向量集，证明 W 是 $\mathbb{R}^4$ 的一个子空间.（利用习题 11 的方法.）
13. 令 $\boldsymbol{v}_1=\begin{bmatrix}1\\0\\-1\end{bmatrix}$, $\boldsymbol{v}_2=\begin{bmatrix}2\\1\\3\end{bmatrix}$, $\boldsymbol{v}_3=\begin{bmatrix}4\\2\\6\end{bmatrix}$, $\boldsymbol{w}=\begin{bmatrix}3\\1\\2\end{bmatrix}$.
   a. $\boldsymbol{w}$ 在 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 吗？$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 中有多少个向量？
   b. $\mathrm{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 中有多少个向量？
   c. $\boldsymbol{w}$ 在由 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 生成的子空间中吗？为什么？
14. 令 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 与 13 题相同，$\boldsymbol{w}=\begin{bmatrix}1\\3\\14\end{bmatrix}$，$\boldsymbol{w}$ 在由 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 生成的子空间中吗？为什么？

在习题 15~18 中，令 W 表示已给形式的所有向量集合，这里 a,b,c 表示任意实数，在每种情形下，要么给出生成 W 的向量集合 S，要么给出一个例子证明 W 不是一个向量空间.

15. $\begin{bmatrix}2a+3b\\-1\\2a-5b\end{bmatrix}$
16. $\begin{bmatrix}1\\3a-5b\\3b+2a\end{bmatrix}$
17. $\begin{bmatrix}2a-b\\3b-c\\3c-a\\3b\end{bmatrix}$
18. $\begin{bmatrix}4a+3b\\0\\a+3b+c\\3b-2c\end{bmatrix}$
19. 一个物体 m 挂在弹簧的末端，若向下拉动物体再放开，这个物体-弹簧系统开始振动，物体从其静止位置的位移 y 由下列形式的函数给出

$$y(t)=c_1\cos\omega t+c_2\sin\omega t \qquad (5)$$

这里 ω 是一个常数，它与弹簧及物体有关.（见下图.）证明由（5）式给出（其中 ω 固定，c_1,c_2 任意）的所有函数的集合是一个向量空间.



20. 定义在 $\mathbb{R}$ 中闭区间 $[a,b]$ 上的全体实连续函数的集合记为 $C[a,b]$，此集合是定义在 $[a,b]$ 上的全体实函数的一个子空间.
    a. 关于连续函数，需要证明什么结论才能说明 $C[a,b]$ 如命题所说的确实是一个子空间？（这些结论在微积分中经常讨论）.
    b. 证明 $\{\boldsymbol{f}\in C[a,b]:\boldsymbol{f}(a)=\boldsymbol{f}(b)\}$ 是 $C[a,b]$ 的一个子空间.

    对固定的正整数 m,n，在通常的矩阵加法运算和实数乘法运算之下，所有 $m\times n$ 矩阵的集合 $M_{m\times n}$ 是一个向量空间.

21. 判定所有形如 $\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$ 的全体矩阵的集合 H 是否是 $M_{2\times2}$ 的一个子空间.
22. 令 F 是一个固定的 3×2 矩阵，令 H 是 $M_{2\times4}$ 中所有使得 $FA=0$（$M_{3\times4}$ 中的零矩阵）的矩阵 A 的全体构成的集合，判断 H 是否是 $M_{2\times4}$ 的一个子空间.

在习题23和习题24中，判断每个命题的真假，给出理由.

23. a. 若 $\boldsymbol{f}$ 是一个向量空间 V 中所有定义在 $\mathbb{R}$ 上的一个实函数，且 $\boldsymbol{f}(t)=0$，则 $\boldsymbol{f}$ 是 V 中的零向量.
    b. 在三维空间中，一个向量是一个箭头.
    c. 对向量空间 V 的一个子集 H，若零向量在 H 中，则 H 是 V 的一个子空间.
    d. 一个子空间仍是一个向量空间.
    e. 模拟信号被用在航天飞机的主控系统，在本章“介绍性实例”中有提到.
24. a. 一个向量是一个向量空间中任意一个元素.
    b. 若 $\boldsymbol{u}$ 是向量空间 V 中一个向量，则 $(-1)\boldsymbol{u}$ 与 $\boldsymbol{u}$ 的负向量相同.
    c. 一个向量空间仍是一个子空间.
    d. $\mathbb{R}^2$ 是 $\mathbb{R}^3$ 的一个子空间.
    e. 一个向量空间 V 的一个子集 H，如果满足以下条件：（i）V 的零向量在 H 中，（ii）$\boldsymbol{u},\boldsymbol{v}$ 和 $\boldsymbol{u}+\boldsymbol{v}$ 在 H 中，（iii）c 是一个标量，$c\boldsymbol{u}$ 在 H 中，则 H 是 V 的一个子空间.

习题 25~29 给出向量空间 V 的公理如何用来证明写在向量空间定义后面的基本性质，用合适的公理序号填空. 由公理 2、4 和 5 可知，对所有 $\boldsymbol{u}$，$\boldsymbol{0}+\boldsymbol{u}=\boldsymbol{u}$ 和 $-\boldsymbol{u}+\boldsymbol{u}=\boldsymbol{0}$ 均成立.

25. 完成下面关于零向量是唯一的证明，假设 $\boldsymbol{w}$ 在 V 中，具有对 V 中所有 $\boldsymbol{u}$，$\boldsymbol{u}+\boldsymbol{w}=\boldsymbol{w}+\boldsymbol{u}=\boldsymbol{u}$ 的性质，特别地，$\boldsymbol{0}+\boldsymbol{w}=\boldsymbol{0}$，但由公理______，$\boldsymbol{0}+\boldsymbol{w}=\boldsymbol{w}$，从而 $\boldsymbol{w}=\boldsymbol{0}+\boldsymbol{w}=\boldsymbol{0}$.
26. 完成下面关于 $-\boldsymbol{u}$ 是 V 中唯一满足 $\boldsymbol{u}+(-\boldsymbol{u})=\boldsymbol{0}$ 的向量的证明. 假设 $\boldsymbol{w}$ 满足 $\boldsymbol{u}+\boldsymbol{w}=\boldsymbol{0}$，用 $-\boldsymbol{u}$ 加到两边，则有

    $(-\boldsymbol{u})+[\boldsymbol{u}+\boldsymbol{w}]=(-\boldsymbol{u})+\boldsymbol{0}$

    $[(-\boldsymbol{u})+\boldsymbol{u}]+\boldsymbol{w}=(-\boldsymbol{u})+\boldsymbol{0}$　这是由公理______（a）

    $\boldsymbol{0}+\boldsymbol{w}=(-\boldsymbol{u})+\boldsymbol{0}$　这是由公理______（b）

    $\boldsymbol{w}=-\boldsymbol{u}$　这是由公理______（c）

27. 在下列证明 $0\boldsymbol{u}=\boldsymbol{0}$ 对任意 V 中的 $\boldsymbol{u}$ 成立的过程中，填写漏掉的公理序号.

    $0\boldsymbol{u}=(0+0)\boldsymbol{u}=0\boldsymbol{u}+0\boldsymbol{u}$　这是由公理______（a）

    两边加 $0\boldsymbol{u}$ 的负向量：

    $0\boldsymbol{u}+(-0\boldsymbol{u})=[0\boldsymbol{u}+0\boldsymbol{u}]+(-0\boldsymbol{u})$

    $0\boldsymbol{u}+(-0\boldsymbol{u})=0\boldsymbol{u}+[0\boldsymbol{u}+(-0\boldsymbol{u})]$

    这是由公理______（b）

    $\boldsymbol{0}=0\boldsymbol{u}+\boldsymbol{0}$　这是由公理______（c）

$\mathbf{0}=0\boldsymbol{u}$　　　　这是由公理______（d）

28. 在下列证明 $c\mathbf{0}=\mathbf{0}$ 对任意数 c 成立的过程中，填写漏掉的公理序号：

$c\mathbf{0}=c(\mathbf{0}+\mathbf{0})$　　　这是由公理______（a）

$=c\mathbf{0}+c\mathbf{0}$　　　这是由公理______（b）

两边加 $c\mathbf{0}$ 的负向量：

$c\mathbf{0}+(-c\mathbf{0})=[c\mathbf{0}+c\mathbf{0}]+(-c\mathbf{0})$

$c\mathbf{0}+(-c\mathbf{0})=c\mathbf{0}+[c\mathbf{0}+(-c\mathbf{0})]$

　　　　这是由公理______（c）

$\mathbf{0}=c\mathbf{0}+\mathbf{0}$　　　这是由公理______（d）

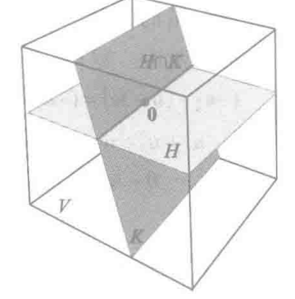
$\mathbf{0}=c\mathbf{0}$　　　这是由公理______（e）

29. 证明 $(-1)\boldsymbol{u}=-\boldsymbol{u}$.（提示：$\boldsymbol{u}+(-1)\boldsymbol{u}=\mathbf{0}$，用习题 27 和习题 26 中某些公理和结论.）

30. 假设对某非零数 $c, c\boldsymbol{u}=\mathbf{0}$, 证明 $\boldsymbol{u}=\mathbf{0}$, 说出你用到的公理和性质.

31. 令 $\boldsymbol{u},\boldsymbol{v}$ 是向量空间 V 中的向量，H 是 V 中任意一个包含 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的子空间，解释为什么 H 也包含 $\mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$，这说明 $\mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$ 是包含 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的 V 中最小子空间.

32. 令 H 和 K 是向量空间 V 的子空间，H 和 K 的交集写成 $H\cap K$，是 V 中既属于 H 又属于 K 的元素的集合，证明 $H\cap K$ 是 V 的一个子空间.（见下图.）在 $\mathbb{R}^2$ 中给出一个例子，证明一般而言，两个子空间的并集不是一个子空间.



33. 给定向量空间 V 的子空间 H 和 K，H 和 K 的和（记为 $H+K$）是 V 中可以写成两个向量（一个向量在 H 中，另一个向量在 K 中）之和的所有向量的集合，即

$$H+K=\{\boldsymbol{w}:\boldsymbol{w}=\boldsymbol{u}+\boldsymbol{v},\text{ 其中 }\boldsymbol{u}\in H,\boldsymbol{v}\in K\}$$

a. 证明 $H+K$ 是 V 的子空间.

b. 证明 H 是 $H+K$ 的子空间，K 是 $H+K$ 的子空间.

34. 假设 $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p$ 和 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_q$ 是向量空间 V 中的向量，并设

$$H=\mathrm{Span}\{\boldsymbol{u}_1,\cdots\boldsymbol{u}_p\},K=\mathrm{Span}\{\boldsymbol{v}_1,\cdots\boldsymbol{v}_q\}$$

证明 $H+K=\mathrm{Span}\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p,\boldsymbol{v}_1,\cdots,\boldsymbol{v}_q\}$.

35. [M] 证明 $\boldsymbol{w}$ 在由 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 生成的 $\mathbb{R}^4$ 的子空间中，这里

$$\boldsymbol{w}=\begin{bmatrix}9\\-4\\-4\\7\end{bmatrix},\ \boldsymbol{v}_1=\begin{bmatrix}8\\-4\\-3\\9\end{bmatrix},\ \boldsymbol{v}_2=\begin{bmatrix}-4\\3\\-2\\-8\end{bmatrix},\ \boldsymbol{v}_3=\begin{bmatrix}-7\\6\\-5\\-18\end{bmatrix}$$

36. [M] 判定 $\boldsymbol{y}$ 是否在由 A 的列生成的 $\mathbb{R}^4$ 的子空间中，这里

$$\boldsymbol{y}=\begin{bmatrix}-4\\-8\\6\\-5\end{bmatrix},\ A=\begin{bmatrix}3&-5&-9\\8&7&-6\\-5&-8&3\\2&-2&-9\end{bmatrix}$$

37. [M] 向量空间 $H=\mathrm{Span}\{1,\cos^2 t,\cos^4 t,\cos^6 t\}$ 至少包含两个有趣的函数，它们在以后的习题中用到：

$$\boldsymbol{f}(t)=1-8\cos^2 t+8\cos^4 t$$
$$\boldsymbol{g}(t)=-1+18\cos^2 t-48\cos^4 t+32\cos^6 t$$

对 $0\leqslant t\leqslant 2\pi$，研究 $\boldsymbol{f}$ 的图像，对 $\boldsymbol{f}(t)$ 推测一个简单的公式，通过绘图比较 $1+\boldsymbol{f}(t)$ 与你的 $\boldsymbol{f}(t)$ 的公式，验证你的猜测.（你将会看到常函数 1.）对 $\boldsymbol{g}$ 重复你的研究.

38. [M] 对下列函数，重复习题 37.

$$\boldsymbol{f}(t)=3\sin t-4\sin^3 t$$
$$\boldsymbol{g}(t)=1-8\sin^2 t+8\sin^4 t$$
$$\boldsymbol{h}(t)=5\sin t-20\sin^3 t+16\sin^5 t$$

它们均在向量空间 $\mathrm{Span}\{1,\sin t,\sin^2 t,\cdots,\sin^5 t\}$ 中.

练习题答案

1. 任取 H 中的 $\boldsymbol{u}$，比如取 $\boldsymbol{u}=\begin{bmatrix}3\\7\end{bmatrix}$，任取 $c\neq 1$，比如取 $c=2$，则 $c\boldsymbol{u}=\begin{bmatrix}6\\14\end{bmatrix}$，如果它在 V 中，则存在某 s 使得

$$\begin{bmatrix} 3s \\ 2+5s \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$$

则 $s=2$ 同时 $s=12/5$，这是不可能的，所以 $2\boldsymbol{u}$ 不在 H 中，H 不是一个向量空间.

2. $\boldsymbol{v}_1 = 1\boldsymbol{v}_1 + 0\boldsymbol{v}_2 + \cdots + 0\boldsymbol{v}_p$，这表明 $\boldsymbol{v}_1$ 是 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p$ 的一个线性组合，所以 $\boldsymbol{v}_1$ 在 W 中，一般 $\boldsymbol{v}_k$ 也在 W 中，这是因为 $\boldsymbol{v}_k = 0\boldsymbol{v}_1 + \cdots + 0\boldsymbol{v}_{k-1} + 1\boldsymbol{v}_k + 0\boldsymbol{v}_{k+1} + \cdots + 0\boldsymbol{v}_p$.

4.2 零空间、列空间和线性变换

在线性代数的应用中，$\mathbb{R}^n$ 的子空间通常由以下两种方式产生：(1) 齐次线性方程组的解集；或 (2) 某些确定向量的线性组合的集合. 本节中，我们比较和构造这两类子空间，体验子空间的概念. 事实上，正如你将要看到的，我们从 1.3 节以来一直与子空间打交道，这里的主要新特征是术语. 本节包括线性变换的值域与核的讨论.

矩阵的零空间

考虑下列齐次方程组

$$\begin{aligned} x_1 - 3x_2 - 2x_3 &= 0 \\ -5x_1 + 9x_2 + x_3 &= 0 \end{aligned} \tag{1}$$

用矩阵的方式，此方程组可写成 $A\boldsymbol{x} = \boldsymbol{0}$，这里

$$A = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix} \tag{2}$$

所有满足 (1) 的 $\boldsymbol{x}$ 的集合称为方程组 (1) 的**解集**，通常将这个解集直接与矩阵 A 和方程 $A\boldsymbol{x}=\boldsymbol{0}$ 联系起来是方便的，我们称满足 $A\boldsymbol{x}=\boldsymbol{0}$ 的所有 $\boldsymbol{x}$ 的集合为矩阵 A 的**零空间**.

定义 $m \times n$ 矩阵 A 的零空间写成 $\mathrm{Nul}\,A$，是齐次方程 $A\boldsymbol{x}=\boldsymbol{0}$ 的全体解的集合，用集合符号表示，即

$$\mathrm{Nul}\,A = \{\boldsymbol{x} : \boldsymbol{x} \in \mathbb{R}^n, A\boldsymbol{x} = \boldsymbol{0}\}$$

$\mathrm{Nul}\,A$ 的更进一步的描述为 $\mathbb{R}^n$ 中在线性变换 $\boldsymbol{x} \mapsto A\boldsymbol{x}$ 下映射到 $\mathbb{R}^m$ 中的零向量的全体向量的集合，见图 4-11.

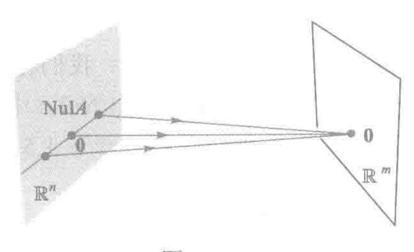


图 4-11

例 1 设 A 为 (2) 式中所示，令 $\boldsymbol{u} = \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix}$，确定 $\boldsymbol{u}$ 是否属于 A 的零空间.

解 为验证 $\boldsymbol{u}$ 是否满足 $A\boldsymbol{u}=\boldsymbol{0}$，简单计算

$$A\boldsymbol{u}=\begin{bmatrix}1 & -3 & -2\\ -5 & 9 & 1\end{bmatrix}\begin{bmatrix}5\\ 3\\ -2\end{bmatrix}=\begin{bmatrix}5-9+4\\ -25+27-2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

所以 $\boldsymbol{u}\in \text{Nul}\,A$. ■

空间这个词用在零空间上是合适的，因为一个矩阵的零空间是一个向量空间，在下列定理中会见到.

定理 2 $m\times n$ 矩阵 A 的零空间是 $\mathbb{R}^n$ 的一个子空间，等价地，m 个方程、n 个未知数的齐次线性方程组 $A\boldsymbol{x}=\boldsymbol{0}$ 的全体解的集合是 $\mathbb{R}^n$ 的一个子空间.

证 由于 A 有 n 个列，$\text{Nul}\,A$ 当然是 $\mathbb{R}^n$ 的一个子集，我们必须证明 $\text{Nul}\,A$ 满足子空间的 3 个性质. $\boldsymbol{0}$ 当然在 $\text{Nul}\,A$ 中，其次令 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 表示 $\text{Nul}\,A$ 中任意两个向量，则

$$A\boldsymbol{u}=\boldsymbol{0},\ A\boldsymbol{v}=\boldsymbol{0}$$

为证 $\boldsymbol{u}+\boldsymbol{v}$ 在 $\text{Nul}\,A$ 中，必须证 $A(\boldsymbol{u}+\boldsymbol{v})=\boldsymbol{0}$. 利用矩阵乘法的性质，有

$$A(\boldsymbol{u}+\boldsymbol{v})=A\boldsymbol{u}+A\boldsymbol{v}=\boldsymbol{0}+\boldsymbol{0}=\boldsymbol{0}$$

从而 $\boldsymbol{u}+\boldsymbol{v}\in \text{Nul}\,A$，于是 $\text{Nul}\,A$ 对向量加法是封闭的. 最后，若 c 是任意一个数，则

$$A(c\boldsymbol{u})=c(A\boldsymbol{u})=c(\boldsymbol{0})=\boldsymbol{0}$$

从而 $c\boldsymbol{u}\in \text{Nul}\,A$，于是 $\text{Nul}\,A$ 是 $\mathbb{R}^n$ 的一个子空间. ■

例 2 令 H 是 $\mathbb{R}^4$ 中坐标 a,b,c,d 满足方程 $a-2b+5c=d$ 且 $c-a=b$ 的所有向量的集合，证明 H 是 $\mathbb{R}^4$ 的一个子空间.

解 通过重新调整描述 H 的元素的方程组，我们发现 H 即下列齐次线性方程组的解集：

$$a-2b+5c-d=0$$
$$-a-b+c=0$$

由定理 2，H 是 $\mathbb{R}^4$ 的一个子空间. ■

确定集合 H 的线性方程组是齐次的这个条件是很重要的，否则，其解集不能确定一个子空间（因为零向量不是非齐次方程组的解），而且在某些情形下，解集可能是空集.

Nul A 的一个显式刻画

$\text{Nul}\,A$ 中的向量与 A 中的数值之间没有明显的关系. 我们称 $\text{Nul}\,A$ 被隐式地定义，这是由于它被一个必须要检验的条件所定义，没有明确地给出 $\text{Nul}\,A$ 中元素. 然而，当我们解出方程 $A\boldsymbol{x}=\boldsymbol{0}$，我们就得到 $\text{Nul}\,A$ 的显式刻画，让我们先复习一下 1.5 节中已有的解题步骤.

例 3 求矩阵 A 的零空间的生成集，其中

$$A=\begin{bmatrix}-3 & 6 & -1 & 1 & -7\\ 1 & -2 & 2 & 3 & -1\\ 2 & -4 & 5 & 8 & -4\end{bmatrix}$$

解 第一步是求 $A\boldsymbol{x}=\boldsymbol{0}$ 的关于自由变量的通解，通过行化简增广矩阵 $[A\ \ \boldsymbol{0}]$ 化为简化阶梯形矩阵为

$$\begin{bmatrix} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{aligned} x_1 - 2x_2 \qquad - x_4 + 3x_5 &= 0 \\ x_3 + 2x_4 - 2x_5 &= 0 \\ 0 &= 0 \end{aligned}$$

通解为 $x_1 = 2x_2 + x_4 - 3x_5, x_3 = -2x_4 + 2x_5, x_2, x_4, x_5$ 是自由变量. 其次，将通解给出的向量分解为向量组合，用自由变量作权，即：

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}$$

$$\uparrow \qquad \uparrow \qquad \uparrow$$
$$\boldsymbol{u} \qquad \boldsymbol{v} \qquad \boldsymbol{w}$$

$$= x_2\boldsymbol{u} + x_4\boldsymbol{v} + x_5\boldsymbol{w} \tag{3}$$

$\boldsymbol{u}, \boldsymbol{v}$ 和 $\boldsymbol{w}$ 的每一个线性组合都是 Nul A 中的一个元素，从而 $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ 是 Nul A 的一个生成集. ■

关于例 3 的解应该得到以下两点事实，它们对所有此类问题均适合，我们将在后面用到.

1. 由例 3 中的方法产生的生成集必然是线性无关的，这是因为自由变量是生成向量上的权. 比如，见（3）式中解向量的第 2，4，5 个数，注意到只有当 x_2, x_4, x_5 全为零时 $x_2\boldsymbol{u} + x_4\boldsymbol{v} + x_5\boldsymbol{w}$ 为零.

2. Nul A 的生成集中向量的个数等于方程 $A\boldsymbol{x} = \boldsymbol{0}$ 中自由变量的个数.

矩阵的列空间

与矩阵相关的另一个重要的子空间是它的列空间. 与零空间不同，列空间可以由向量的线性组合显式定义.

定义 *$m \times n$ 矩阵的列空间（记为 Col A）是由 A 的列的所有线性组合组成的集合，若 $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$，则 Col $A = \text{Span}\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n\}$.*

由于 Col $A = \text{Span}\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n\}$ 是一个子空间，由定理 1 及 Col A 的定义和 A 的列在 $\mathbb{R}^m$ 中这一事实得到以下定理.

定理 3 *$m \times n$ 矩阵 A 的列空间是 $\mathbb{R}^m$ 的一个子空间.*

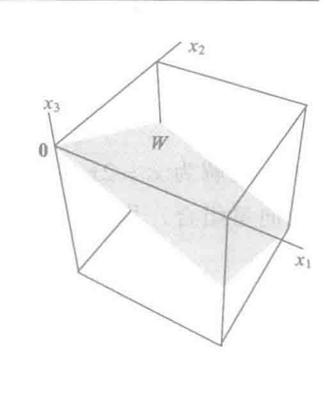
注意到 Col A 中一个典型向量可写成 $A\boldsymbol{x}$ 的形式，其中 $\boldsymbol{x}$ 为某向量，因为记号 $A\boldsymbol{x}$ 表示 A 的列向量的一个线性组合，即

$$\text{Col } A = \{\boldsymbol{b} : \boldsymbol{b} = A\boldsymbol{x}, \boldsymbol{x} \in \mathbb{R}^n\}$$

代表 Col A 中向量的记号 $A\boldsymbol{x}$ 也表明 Col A 是线性变换 $\boldsymbol{x} \mapsto A\boldsymbol{x}$ 的值域，我们将在本节的最后讨论这个问题.

例 4 求一个矩阵 A，使得 $W = \text{Col } A$.

$$W=\left\{\begin{bmatrix}6a-b\\a+b\\-7a\end{bmatrix}:a,b\in\mathbb{R}\right\}$$



解 首先将 W 写成线性组合的集合

$$W=\left\{a\begin{bmatrix}6\\1\\-7\end{bmatrix}+b\begin{bmatrix}-1\\1\\0\end{bmatrix}:a,b\in\mathbb{R}\right\}=\text{Span}\left\{\begin{bmatrix}6\\1\\-7\end{bmatrix},\begin{bmatrix}-1\\1\\0\end{bmatrix}\right\}$$

其次，用生成集中的向量作为 A 的列，令 $A=\begin{bmatrix}6&-1\\1&1\\-7&0\end{bmatrix}$，则 $W=\text{Col}\,A$，A 为所求矩阵. ■

回顾 1.4 节中的定理 4，A 的列生成 $\mathbb{R}^m$ 当且仅当方程 $A\mathbf{x}=\mathbf{b}$ 对任一个 $\mathbf{b}$ 有解，我们可以按下列方式重述这一事实.

$m\times n$ 矩阵 A 的列空间等于 $\mathbb{R}^m$ 当且仅当方程 $A\mathbf{x}=\mathbf{b}$ 对 $\mathbb{R}^m$ 中每个 $\mathbf{b}$ 有一个解.

Nul*A* 与 Col*A* 之间的对比

一个矩阵的零空间和列空间之间的关系如何是一个自然的问题，事实上，这两个空间是很不一样的，从例 5~7 中将看到这一点. 然而，这两个空间之间的一个令人感到意外的关联将出现在 4.6 节中，这需要用到更多的理论.

例 5 令 $A=\begin{bmatrix}2&4&-2&1\\-2&-5&7&3\\3&7&-8&6\end{bmatrix}$.

a. 若 A 的列空间是 $\mathbb{R}^k$ 的一个子空间，k 是多少？

b. 若 A 的零空间是 $\mathbb{R}^k$ 的一个子空间，k 是多少？

解 a. A 的每一列含有 3 个数，所以 $\text{Col}\,A$ 是 $\mathbb{R}^3$ 的一个子空间.

b. 使得 $A\mathbf{x}$ 有定义的一个向量 $\mathbf{x}$ 必须有 4 个数. 所以 $\text{Nul}\,A$ 是 $\mathbb{R}^4$ 的一个子空间. ■

当一个矩阵不是方阵时，比如例 5 中的矩阵，$\text{Nul}\,A$ 中的向量与 $\text{Col}\,A$ 中的向量分别在完全不同的“域”中. 例如，我们已讨论过 $\mathbb{R}^3$ 中向量的线性组合不能产生 $\mathbb{R}^4$ 中的一个向量. 当 A 为方阵时，$\text{Nul}\,A$ 和 $\text{Col}\,A$ 确实都具有零向量，并且在特殊情形下它们可能具有一些相同的非零向量.

例 6 对例 5 中的 A，分别找出 $\text{Col}\,A$ 和 $\text{Nul}\,A$ 中的一个非零向量.

解 容易找到 $\text{Col}\,A$ 中的一个向量，A 的任一个列都可以，比如 $\begin{bmatrix}2\\-2\\3\end{bmatrix}$，为了找到 $\text{Nul}\,A$ 中的一个非零向量，我们已作了一些工作，将增广矩阵 $[A\ \ \mathbf{0}]$ 行化简，得

$$[A\ \ \mathbf{0}]\sim\begin{bmatrix}1&0&9&0&0\\0&1&-5&0&0\\0&0&0&1&0\end{bmatrix}$$

从而若 $\boldsymbol{x}$ 满足 $A\boldsymbol{x}=\boldsymbol{0}$，则 $x_1=-9x_3, x_2=5x_3, x_4=0, x_3$ 是自由变量，x_3 取一个非零值，比如 $x_3=1$，我们得到 Nul A 中一个向量，即 $\boldsymbol{x}=(-9,5,1,0)$．■

例 7 对例 5 中的 A，令 $\boldsymbol{u}=\begin{bmatrix}3\\-2\\-1\\0\end{bmatrix}, \boldsymbol{v}=\begin{bmatrix}3\\-1\\3\end{bmatrix}$.

a. 判定 $\boldsymbol{u}$ 是否在 Nul A 中，$\boldsymbol{u}$ 是否在 Col A 中.

b. 判定 $\boldsymbol{v}$ 是否在 Col A 中，$\boldsymbol{v}$ 是否在 Nul A 中.

解 a. 这里不需要 Nul A 的一个显式刻画，简单地计算乘积 $A\boldsymbol{u}$.

$$A\boldsymbol{u}=\begin{bmatrix}2&4&-2&1\\-2&-5&7&3\\3&7&-8&6\end{bmatrix}\begin{bmatrix}3\\-2\\-1\\0\end{bmatrix}=\begin{bmatrix}0\\-3\\3\end{bmatrix}\neq\begin{bmatrix}0\\0\\0\end{bmatrix}$$

显然，$\boldsymbol{u}$ 不是 $A\boldsymbol{x}=\boldsymbol{0}$ 的一个解，所以 $\boldsymbol{u}$ 不在 Nul A 中，由于 Col A 是 $\mathbb{R}^3$ 的一个子空间，$\boldsymbol{u}$ 具有 4 个数值，$\boldsymbol{u}$ 不可能在 Col A 中.

b. 将 $[A \ \ \boldsymbol{v}]$ 简化成阶梯形

$$[A \ \ \boldsymbol{v}]=\begin{bmatrix}2&4&-2&1&3\\-2&-5&7&3&-1\\3&7&-8&6&3\end{bmatrix}\sim\begin{bmatrix}2&4&-2&1&3\\0&1&-5&-4&-2\\0&0&0&17&1\end{bmatrix}$$

由此可见，方程 $A\boldsymbol{x}=\boldsymbol{v}$ 是相容的，所以 $\boldsymbol{v}$ 在 Col A 中，由于 Nul A 是 $\mathbb{R}^4$ 的子空间，$\boldsymbol{v}$ 具有 3 个数值，$\boldsymbol{v}$ 不可能在 Nul A 中.■

表 4-1 可以作为我们已经学过的关于 Nul A 与 Col A 的一个总结. 其中的第 8 条是 1.9 节中定理 11 和定理 12（a）的另一种形式.

表 4-1 对 $m\times n$ 矩阵 A，Nul A 与 Col A 之间的对比

Nul A	Col A
1. Nul A 是 $\mathbb{R}^n$ 的一个子空间.	1. Col A 是 $\mathbb{R}^m$ 的一个子空间.
2. Nul A 是隐式定义的，即仅给出了一个 Nul A 中向量必须满足的条件 ($A\boldsymbol{x}=\boldsymbol{0}$)．	2. Col A 是显式定义，即明确指出如何建立 Col A 中的向量.
3. 求 Nul A 中的向量需要时间，需要对 $[A \ \ \boldsymbol{0}]$ 做行变换.	3. 容易求出 Col A 中的向量，A 中的列就是 Col A 中的向量，其余的可由 A 的列表示出来.
4. Nul A 与 A 的数值之间没有明显的关系.	4. Col A 与 A 的数值之间有明显的关系，因为 A 的列就在 Col A 中.
5. Nul A 中的一个典型向量 $\boldsymbol{v}$ 具有 $A\boldsymbol{v}=\boldsymbol{0}$ 的性质.	5. Col A 中一个典型向量 $\boldsymbol{v}$ 具有方程 $A\boldsymbol{x}=\boldsymbol{v}$ 是相容的性质.
6. 给一个特定的向量 $\boldsymbol{v}$，容易判断 $\boldsymbol{v}$ 是否在 Nul A 中，仅需计算 $A\boldsymbol{v}$．	6. 给一个特定的向量 $\boldsymbol{v}$，弄清 $\boldsymbol{v}$ 是否在 Col A 中需要时间，需要对 $[A \ \ \boldsymbol{v}]$ 作行变换.
7. Nul $A=\{\boldsymbol{0}\}$ 当且仅当方程 $A\boldsymbol{x}=\boldsymbol{0}$ 仅有一个平凡解.	7. Col $A=\mathbb{R}^m$ 当且仅当方程 $A\boldsymbol{x}=\boldsymbol{b}$ 对每一个 $\boldsymbol{b}\in\mathbb{R}^m$ 有一个解.
8. Nul $A=\{\boldsymbol{0}\}$ 当且仅当线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 是一对一的.	8. Col $A=\mathbb{R}^m$ 当且仅当线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 将 $\mathbb{R}^n$ 映上到 $\mathbb{R}^m$．

线性变换的核与值域

我们经常需要用线性变换而不是矩阵来描述除 $\mathbb{R}^n$ 以外的向量空间的子空间，为了更准确，将 1.8 节中给出的定义进行推广.

定义 由向量空间 V 映射到向量空间 W 内的**线性变换** T 是一个规则，它将 V 中每个向量 $\boldsymbol{x}$ 映射成 W 中唯一向量 $T(\boldsymbol{x})$，且满足：

（i）$T(\boldsymbol{u}+\boldsymbol{v})=T(\boldsymbol{u})+T(\boldsymbol{v})$，对 V 中所有 $\boldsymbol{u},\boldsymbol{v}$ 均成立.

（ii）$T(c\boldsymbol{u})=cT(\boldsymbol{u})$，对 V 中所有 $\boldsymbol{u}$ 及所有数 c 均成立.

线性变换 T 的**核**（或**零空间**）是 V 中所有满足 $T(\boldsymbol{u})=\boldsymbol{0}$ 的向量 $\boldsymbol{u}$ 的集合（$\boldsymbol{0}$ 为 W 中的零向量）. T 的**值域**是 W 中所有具有形式 $T(\boldsymbol{x})$（任意 $\boldsymbol{x}\in V$）的向量的集合. 如果 T 是由一个矩阵变换得到的，比如对某矩阵 A，$T(\boldsymbol{x})=A\boldsymbol{x}$，则 T 的核与值域恰好是前面定义的 A 的零空间和列空间.

不难证明 T 的核是 V 的一个子空间. 证明在本质上与定理 2 相同，T 的值域也是 W 的一个子空间.见图 4-12 和习题 30.

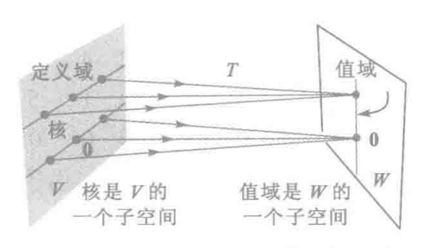


图 4-12 与线性变换相关的子空间

在应用中，一个子空间往往由一个适当的线性变换的核或值域产生，比如，一个齐次线性微分方程的全部解的集合被判明是一个线性变换的核. 典型地，这样一个线性变换用关于一个函数的一阶或高阶导数描述. 要解释这个结论将使我们远离主题，所以我们仅给出两个例子，第一个例子解释微分运算为什么是一个线性变换.

例 8（需要微积分的知识） 令 V 是定义在区间 $[a,b]$ 上的所有连续可导的实函数构成的向量空间，令 W 是 $[a,b]$ 上所有连续函数构成的向量空间，且令 $D:V\to W$ 是将 V 中 f 变为其导数 f' 的变换，由微积分中两个简单的微分法则有

$$D(f+g)=D(f)+D(g)\ ,\qquad D(cf)=cD(f)$$

于是，D 是一个线性变换，可以证明 D 的核是 $[a,b]$ 上的常函数的集合，D 的值域是 $[a,b]$ 上所有连续函数的集合 W . ■

例 9（需要微积分的知识） 微分方程

$$y''+\omega^2 y=0 \tag{4}$$

这里 ω 是常数，常常用来描述物理系统的一个变化过程，比如负重弹簧的振动，摆的运动，以及电感-电容电路中的电压.（4）的解集恰好是将函数 $y=f(t)$ 映成函数 $f''(t)+\omega^2 f(t)$ 的线性变换

的核. 寻找这个向量空间的一个显式刻画是微分方程方面的一个问题. 其解集结果是 4.1 节习题 19 中所说的空间. ■

练习题

1. 令 $W=\left\{\begin{bmatrix} a \\ b \\ c \end{bmatrix}: a-3b-c=0\right\}$，用两种不同的方法证明 W 是 $\mathbb{R}^3$ 的一个子空间.（用两个定理.）

2. 令 $A=\begin{bmatrix} 7 & -3 & 5 \\ -4 & 1 & -5 \\ -5 & 2 & -4 \end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$，$\boldsymbol{w}=\begin{bmatrix} 7 \\ 6 \\ -3 \end{bmatrix}$，假设已知方程 $A\boldsymbol{x}=\boldsymbol{v}$ 和 $A\boldsymbol{x}=\boldsymbol{w}$ 都是相容的，关于方程 $A\boldsymbol{x}=\boldsymbol{v}+\boldsymbol{w}$，你能得出什么结果？

习题 4.2

1. 判定 $\boldsymbol{w}=\begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix}$ 是否在 Nul A 中，这里

$$A=\begin{bmatrix} 3 & -5 & -3 \\ 6 & -2 & 0 \\ -8 & 4 & 1 \end{bmatrix}$$

2. 判定 $\boldsymbol{w}=\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ 是否在 Nul A 中，这里

$$A=\begin{bmatrix} 2 & 6 & 4 \\ -3 & 2 & 5 \\ -5 & -4 & 1 \end{bmatrix}$$

在习题 3~6 中，通过求出张成零空间的向量，求出 Nul A 的一个显式表示.

3. $A=\begin{bmatrix} 1 & 2 & 4 & 0 \\ 0 & 1 & 3 & -2 \end{bmatrix}$

4. $A=\begin{bmatrix} 1 & -3 & 2 & 0 \\ 0 & 0 & 3 & 0 \end{bmatrix}$

5. $A=\begin{bmatrix} 1 & -4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -5 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$

6. $A=\begin{bmatrix} 1 & 3 & -4 & -3 & 1 \\ 0 & 1 & -3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

在习题 7~14 中，或者利用一个适当的定理证明给出的集合 W 是一个向量空间，或者举例说明它不是一个向量空间.

7. $\left\{\begin{bmatrix} a \\ b \\ c \end{bmatrix}: a+b+c=2\right\}$

8. $\left\{\begin{bmatrix} r \\ s \\ t \end{bmatrix}: 3r-2=3s+t\right\}$

9. $\left\{\begin{bmatrix} p \\ q \\ r \\ s \end{bmatrix}: \begin{matrix} p-2q=4s \\ 2p=s+5r \end{matrix}\right\}$

10. $\left\{\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}: \begin{matrix} 3a+b=c \\ a+b+2c=2d \end{matrix}\right\}$

11. $\left\{\begin{bmatrix} s-2t \\ 3+3s \\ 3s+3t \\ 2s \end{bmatrix}: s,t \text{ 为实数}\right\}$

12. $\left\{\begin{bmatrix} 3p-5q \\ 4q \\ p \\ q+1 \end{bmatrix}: p,q \text{ 为实数}\right\}$

13. $\left\{\begin{bmatrix} c-6d \\ d \\ c \end{bmatrix}: c,d \text{ 为实数}\right\}$

14. $\left\{\begin{bmatrix}-s+3t\\s-2t\\5s-t\end{bmatrix}: s,t \text{ 为实数}\right\}$

在习题 15 和习题 16 中，求 A 使得给出的集合为 $\text{Col}\,A$.

15. $\left\{\begin{bmatrix}2s+t\\r-s+2t\\3r+s\\2r-s-t\end{bmatrix}: r,s,t \text{ 为实数}\right\}$

16. $\left\{\begin{bmatrix}b-c\\2b+3d\\b+3c-3d\\c+d\end{bmatrix}: b,c,d \text{为实数}\right\}$

对习题 17~20 中的矩阵，(a) 求 k 使得 $\text{Nul}\,A$ 是 $\mathbb{R}^k$ 的一个子空间，(b) 求 k 使得 $\text{Col}\,A$ 是 $\mathbb{R}^k$ 的一个子空间.

17. $A=\begin{bmatrix}6&-4\\-3&2\\-9&6\\9&-6\end{bmatrix}$

18. $A=\begin{bmatrix}5&-2&3\\-1&0&-1\\0&-2&-2\\-5&7&2\end{bmatrix}$

19. $A=\begin{bmatrix}4&5&-2&6&0\\1&1&0&1&0\end{bmatrix}$

20. $A=\begin{bmatrix}1&-3&9&0&-5\end{bmatrix}$

21. 对习题 17 中的 A，分别求 $\text{Nul}\,A$ 和 $\text{Col}\,A$ 中的一个非零向量.

22. 对习题 3 中的 A，分别求 $\text{Nul}\,A$ 和 $\text{Col}\,A$ 中的一个非零向量.

23. 令 $A=\begin{bmatrix}-2&4\\-1&2\end{bmatrix}, \boldsymbol{w}=\begin{bmatrix}2\\1\end{bmatrix}$，判断 $\boldsymbol{w}$ 是否在 $\text{Col}\,A$ 中，$\boldsymbol{w}$ 是否在 $\text{Nul}\,A$ 中.

24. 令 $A=\begin{bmatrix}10&-8&-2&-2\\0&2&2&-2\\1&-1&6&0\\1&1&0&-2\end{bmatrix}, \boldsymbol{w}=\begin{bmatrix}2\\2\\0\\2\end{bmatrix}$，判定 $\boldsymbol{w}$ 是否在 $\text{Col}\,A$ 中，$\boldsymbol{w}$ 是否在 $\text{Nul}\,A$ 中.

在习题 25 和习题 26 中，A 表示一个 $m\times n$ 矩阵. 对每个命题判断真假，给出理由.

25. a. A 的零空间是方程 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集.
    b. $m\times n$ 矩阵的零空间包含在 $\mathbb{R}^m$ 中.
    c. A 的列空间是映射 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的值域.
    d. 若方程 $A\boldsymbol{x}=\boldsymbol{b}$ 是相容的，则 $\text{Col}\,A$ 就是 $\mathbb{R}^m$.
    e. 一个线性变换的核是一个向量空间.
    f. $\text{Col}\,A$ 是对某 $\boldsymbol{x}$ 所有能写成 $A\boldsymbol{x}$ 的向量的集合.

26. a. 零空间是向量空间.
    b. $m\times n$ 矩阵的列空间包含在 $\mathbb{R}^m$ 中.
    c. $\text{Col}\,A$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的所有解的集合.
    d. $\text{Nul}\,A$ 是映射 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的核.
    e. 一个线性变换的值域是一个向量空间.
    f. 一个齐次线性微分方程的所有解的集合是一个线性变换的核.

27. 可以证明下列方程组的一组解为 $x_1=3, x_2=2, x_3=-1$. 利用本节中的理论解释为什么 $x_1=30, x_2=20, x_3=-10$ 是另一组解（观察这两组解之间的关系，不要作另外的计算）.

$$\begin{aligned}x_1-3x_2-3x_3&=0\\-2x_1+4x_2+2x_3&=0\\-x_1+5x_2+7x_3&=0\end{aligned}$$

28. 考察下列两个方程组：

$$\begin{aligned}5x_1+\ \ x_2-3x_3&=0\\-9x_1+2x_2+5x_3&=1\\4x_1+\ \ x_2-6x_3&=9\end{aligned}\qquad\begin{aligned}5x_1+\ \ x_2-3x_3&=0\\-9x_1+2x_2+5x_3&=5\\4x_1+\ \ x_2-6x_3&=45\end{aligned}$$

可以证明第一个方程组有解，利用这个事实和本节中的理论解释为什么第二个方程组一定也有解.（不要作行变换.）

29. 定理 3 的证明如下：给一个 $m\times n$ 矩阵 A，$\text{Col}\,A$ 中任一元素均为形如 $A\boldsymbol{x}, \boldsymbol{x}\in\mathbb{R}^n$ 的形式. 令 $A\boldsymbol{x}$ 和 $A\boldsymbol{w}$ 分别为 $\text{Col}\,A$ 中任意两个向量.
    a. 解释为什么零向量在 $\text{Col}\,A$ 中.
    b. 证明向量 $A\boldsymbol{x}+A\boldsymbol{w}$ 在 $\text{Col}\,A$ 中.
    c. 给一数 c，证明 $c(A\boldsymbol{x})$ 在 $\text{Col}\,A$ 中.

30. 令 $T:V \to W$ 是一个从向量空间 V 到向量空间 W 中的线性变换，证明 T 的值域是 W 的一个子空间．（提示：值域中的典型元素具有形式 $T(\boldsymbol{x})$ 和 $T(\boldsymbol{w})$，其中 $\boldsymbol{x},\boldsymbol{w}$ 属于 V．）

31. 定义 $T:\mathbb{P}_2 \to \mathbb{R}^2, T(\boldsymbol{p})=\begin{bmatrix} \boldsymbol{p}(0) \\ \boldsymbol{p}(1) \end{bmatrix}$，比如，若 $\boldsymbol{p}(t)=3+5t+7t^2$，则 $T(\boldsymbol{p})=\begin{bmatrix} 3 \\ 15 \end{bmatrix}$．

   a. 证明 T 是一个线性变换．（提示：对 $\mathbb{P}_2$ 中的任意多项式 $\boldsymbol{p},\boldsymbol{q}$，计算 $T(\boldsymbol{p}+\boldsymbol{q})$ 和 $T(c\boldsymbol{p})$．）

   b. 求 $\mathbb{P}_2$ 中的一个多项式 $\boldsymbol{p}$ 使之生成 T 的核并刻画 T 的值域．

32. 由 $T(\boldsymbol{p})=\begin{bmatrix} \boldsymbol{p}(0) \\ \boldsymbol{p}(0) \end{bmatrix}$ 定义一个线性变换 $T:\mathbb{P}_2 \to \mathbb{R}^2$ 求 $\mathbb{P}_2$ 中多项式 $\boldsymbol{p}_1$ 和 $\boldsymbol{p}_2$，使之生成 T 的核并刻画 T 的值域．

33. 令 $M_{2\times2}$ 是所有 2×2 矩阵组成的向量空间，定义 $T:M_{2\times2} \to M_{2\times2}, T(A)=A+A^{\mathrm{T}}$，这里

$$A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

   a. 证明 T 是一个线性变换．

   b. 设 B 是 $M_{2\times2}$ 中任一个满足 $B^{\mathrm{T}}=B$ 的矩阵，求 $M_{2\times2}$ 中的矩阵 A 使得 $T(A)=B$．

   c. 证明 T 的值域是 $M_{2\times2}$ 中满足性质 $B^{\mathrm{T}}=B$ 的 B 的集合．

   d. 给出 T 的核．

34. （需要微积分的知识）定义 $T:C[0,1] \to C[0,1]$ 如下：对 $\boldsymbol{f} \in C[0,1]$，令 $T(\boldsymbol{f})$ 是 $\boldsymbol{f}$ 的满足 $\boldsymbol{F}(0)=0$ 的原函数．证明 T 是一个线性变换并刻画 T 的核．（见 4.1 节习题 20 的注解．）

35. 令 V 和 W 为向量空间，令 $T:V \to W$ 是一个线性变换，给 V 的一子空间 U，令 $T(U)$ 表示所有形如 $T(\boldsymbol{x})$ 的像的集合，这里 $\boldsymbol{x}$ 在 U 中，证明 $T(U)$ 是 W 的一个子空间．

36. 已知 $T:V \to W$ 与 35 题中相同，给 W 的一子空间 Z，令 U 是 V 中所有使得 $T(\boldsymbol{x})$ 在 Z 中的 $\boldsymbol{x}$ 的集合．证明 U 是 V 的一个子空间．

37. [M]判定 $\boldsymbol{w}$ 是否在 A 的列空间中，是否在 A 的零空间中，或同时在两个空间中，这里

$$\boldsymbol{w}=\begin{bmatrix} 1 \\ 1 \\ -1 \\ -3 \end{bmatrix}, A=\begin{bmatrix} 7 & 6 & -4 & 1 \\ -5 & -1 & 0 & -2 \\ 9 & -11 & 7 & -3 \\ 19 & -9 & 7 & 1 \end{bmatrix}$$

38. [M]判定 $\boldsymbol{w}$ 是否在 A 的列空间中，是否在 A 的零空间中，或同时在两个空间中，这里

$$\boldsymbol{w}=\begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix}, A=\begin{bmatrix} -8 & 5 & -2 & 0 \\ -5 & 2 & 1 & -2 \\ 10 & -8 & 6 & -3 \\ 3 & -2 & 1 & 0 \end{bmatrix}$$

39. [M]令 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_5$ 表示矩阵 A 的列，这里

$$A=\begin{bmatrix} 5 & 1 & 2 & 2 & 0 \\ 3 & 3 & 2 & -1 & -12 \\ 8 & 4 & 4 & -5 & 12 \\ 2 & 1 & 1 & 0 & -2 \end{bmatrix}, B=[\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \quad \boldsymbol{a}_4]$$

   a. 解释为什么 $\boldsymbol{a}_3$和$\boldsymbol{a}_5$ 在 B 的列空间中．

   b. 求生成 Nul A 的向量的集合．

   c. 令 $T:\mathbb{R}^5 \to \mathbb{R}^4$ 定义为 $T(\boldsymbol{x})=A\boldsymbol{x}$，解释为什么 T 既不是一一的又不是到上的．

40. [M]令 $H=\mathrm{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}, K=\mathrm{Span}\{\boldsymbol{v}_3,\boldsymbol{v}_4\}$，其中

$$\boldsymbol{v}_1=\begin{bmatrix} 5 \\ 3 \\ 8 \end{bmatrix}, \boldsymbol{v}_2=\begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}, \boldsymbol{v}_3=\begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}, \boldsymbol{v}_4=\begin{bmatrix} 0 \\ -12 \\ -28 \end{bmatrix}$$

   则 H 和 K 是 $\mathbb{R}^3$ 的子空间．事实上，H 和 K 是 $\mathbb{R}^3$ 中通过原点的平面，二者相交于一条过原点的直线，求一个非零向量 $\boldsymbol{w}$，使之生成这条直线．（提示：$\boldsymbol{w}$ 可写成 $c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2$，也可写成 $c_3\boldsymbol{v}_3+c_4\boldsymbol{v}_4$，为求 $\boldsymbol{w}$，对未知的 c_j $(j=1,2,3,4)$，解方程 $c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2=c_3\boldsymbol{v}_3+c_4\boldsymbol{v}_4$．）

练习题答案

1. 方法 1：因为 W 是一个齐次线性方程组的全部解的集合，由定理 2，W 是 $\mathbb{R}^3$ 的一子空间（这里方程组仅有一个方程）. 等价地，W 是 1×3 矩阵 $A=[1\quad -3\quad -1]$ 的零空间.

   方法 2：以 b,c 为自由变量解方程 $a-3b-c=0$，任意解具有形式 $\begin{bmatrix}3b+c\\b\\c\end{bmatrix}$，这里 b,c 是任意常数，且

$$\begin{bmatrix}3b+c\\b\\c\end{bmatrix}=b\underset{\uparrow\atop \boldsymbol{v}_1}{\begin{bmatrix}3\\1\\0\end{bmatrix}}+c\underset{\uparrow\atop \boldsymbol{v}_2}{\begin{bmatrix}1\\0\\1\end{bmatrix}}$$

   计算表明 $W=\mathrm{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$，从而由定理 1 知 W 是 $\mathbb{R}^3$ 的一个子空间，我们还可以 a,c 或 a,b 为自由变量解方程 $a-3b-c=0$，得到 W 作为两个向量的线性组合的集合的不同刻画.

2. $\boldsymbol{v}$ 和 $\boldsymbol{w}$ 都在 $\mathrm{Col}\,A$ 中，因 $\mathrm{Col}\,A$ 是一个向量空间，$\boldsymbol{v}+\boldsymbol{w}$ 一定在 $\mathrm{Col}\,A$ 中，即方程 $A\boldsymbol{x}=\boldsymbol{v}+\boldsymbol{w}$ 是相容的.

4.3 线性无关集和基

本节我们确认并研究尽可能“有效率地”生成一个向量空间 V 或一个子空间 H 的子集，关键的思想是线性无关，与定义在 $\mathbb{R}^n$ 中的线性无关一样.

V 中向量的一个指标集 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 称为是**线性无关**的，如果向量方程

$$c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_p\boldsymbol{v}_p=\boldsymbol{0} \tag{1}$$

只有平凡解，即 $c_1=0,\cdots,c_p=0$.[㊀]

集合 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 称为**线性相关**，如果（1）有一个非平凡的解，即存在某些权 $c_1,\cdots,c_p$ 不全为零，使得（1）式成立，此时（1）式称为 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 之间的一个**线性相关关系**.

与 $\mathbb{R}^n$ 中一样，一个仅含一个向量 $\boldsymbol{v}$ 的集是线性无关的当且仅当 $\boldsymbol{v}\neq\boldsymbol{0}$；一个仅含两个向量的集合是线性相关的当且仅当其中一个向量是另一个的倍数；任何含有零向量的集合是线性相关的. 下列定理与 1.7 节中定理 7 证法相同.

定理 4 不少于两个有编号的向量的集合 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$，如果有 $\boldsymbol{v}_1\neq\boldsymbol{0}$，则 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 是线性相关的，当且仅当某 $\boldsymbol{v}_j(j>1)$ 是其前面向量 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_{j-1}$ 的线性组合.

一般向量空间中的线性相关与 $\mathbb{R}^n$ 中的线性相关的主要不同点在于当向量不是 n 元数组时，齐次方程（1）通常不能被写成一个 n 元线性方程组. 换句话说，为了研究方程 $A\boldsymbol{x}=\boldsymbol{0}$，向量不能从一个矩阵 A 的列中得到，取而代之的是我们必须要依靠线性相关的定义和定理 4.

例 1 令 $\boldsymbol{p}_1(t)=1,\boldsymbol{p}_2(t)=t,\boldsymbol{p}_3(t)=4-t$, 则由于 $\boldsymbol{p}_3=4\boldsymbol{p}_1-\boldsymbol{p}_2$，从而 $\{\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3\}$ 是线性相关的. ■

㊀ 在（1）中用 $c_1,\cdots,c_p$ 表示标量而代替如第 1 章中用过的 $x_1,\cdots,x_p$ 是方便的.

例 2 集合$\{\sin t, \cos t\}$在$C[0,1]$中是线性无关的，这是因为作为$C[0,1]$中的向量，$\sin t$和$\cos t$中任一个均不是另一个的倍数. 即不存在数c使得$\cos t = c \cdot \sin t$对任意$t \in [0,1]$均成立（见$\sin t$和$\cos t$的图像）. 然而，$\{\sin t \cos t, \sin 2t\}$是线性相关的，这是因为有$\sin 2t = 2\sin t \cos t$，对任意$t \in [0,1]$均成立. ■

定义 令H是向量空间V的一个子空间，V中向量的指标集$\mathcal{B} = \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_p\}$称为$H$的一个**基**，如果

（i）$\mathcal{B}$是一线性无关集.

（ii）由$\mathcal{B}$生成的子空间与H相同，即$H = \text{Span}\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_p\}$.

因为任一个向量空间都是其自身的子空间，所以基的定义也可用在$H = V$的情形，因而V的一个基是生成V的一个线性无关集. 注意到当$H \neq V$时，条件（ii）蕴涵$\boldsymbol{b}_1, \cdots, \boldsymbol{b}_p$中每个向量都属于$H$，因为正如 4.1 节中所见，$\text{Span}\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_p\}$包含$\boldsymbol{b}_1, \cdots, \boldsymbol{b}_p$.

例 3 令A是一个可逆的$n \times n$矩阵，即$A = [\boldsymbol{a}_1 \ \cdots \ \boldsymbol{a}_n]$，则由可逆矩阵定理，$A$的列组成$\mathbb{R}^n$的一个基，这是因为它们是线性无关的且它们可以生成$\mathbb{R}^n$. ■

例 4 令$\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$是$n \times n$单位矩阵$I_n$的列，即

$$\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \cdots, \boldsymbol{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

集合$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$称为$\mathbb{R}^n$的标准基（图 4-13）.

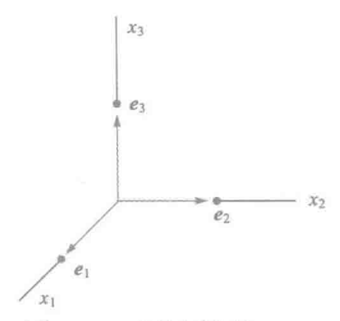


图 4-13 $\mathbb{R}^3$的标准基 ■

例 5 令$\boldsymbol{v}_1 = \begin{bmatrix} 3 \\ 0 \\ -6 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} -4 \\ 1 \\ 7 \end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix}$，判断$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$是否是$\mathbb{R}^3$的一个基.

解 因$\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$是$\mathbb{R}^3$中的 3 个向量，我们可以用几种方法判定矩阵$A = [\boldsymbol{v}_1 \ \boldsymbol{v}_2 \ \boldsymbol{v}_3]$是否可逆，比如通过简单计算得$\det A = 6 \neq 0$，从而$A$可逆. 与例 3 相同，$A$的列是$\mathbb{R}^3$的一个基. ■

例 6 令$S = \{1, t, t^2, \cdots, t^n\}$，证明$S$是$\mathbb{P}_n$的一个基，此基称为$\mathbb{P}_n$的标准基.

解 显然S生成$\mathbb{P}_n$，为证S是线性无关的，假设$c_0, \cdots, c_n$满足

$$c_0 \cdot 1 + c_1 t + c_2 t^2 + \cdots + c_n t^n = \mathbf{0}(t) \tag{2}$$

此式表明左边的多项式与右边的零多项式具有相同的值，由代数的基本定理知，$\mathbb{P}_n$ 中的多项式若有多于 n 个的根则此多项式一定为零多项式，即（2）对任意 t 只有当 $c_0 = \cdots = c_n = 0$ 时成立，于是证明 S 是线性无关的且是 $\mathbb{P}_n$ 的一个基，见图 4-14.

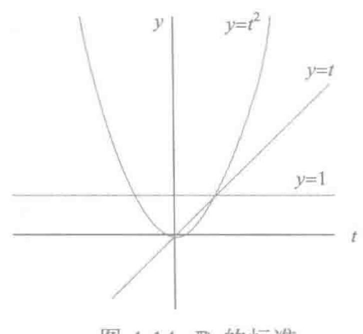


图 4-14 $\mathbb{P}_2$ 的标准 ■

涉及 $\mathbb{P}_n$ 中线性无关和生成的问题，用 4.4 节中讨论的技巧处理最合适.

生成集定理

正如将要看到的，一个基是一个不包含不必要的向量的“高效率”的生成集. 事实上，一个基可以通过由一个生成集中去掉不需要的向量构造出来.

例 7 令 $\boldsymbol{v}_1 = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix} 6 \\ 16 \\ -5 \end{bmatrix}, H = \mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$，注意到 $\boldsymbol{v}_3 = 5\boldsymbol{v}_1 + 3\boldsymbol{v}_2$，证明 $\mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\} = \mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$，然后求子空间 H 的一个基.

解 因为 $c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + 0\boldsymbol{v}_3$，所以 $\mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ 中每个向量都在 H 中，现令 $\boldsymbol{x}$ 为 H 中任一向量，即 $\boldsymbol{x} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_3\boldsymbol{v}_3$，因 $\boldsymbol{v}_3 = 5\boldsymbol{v}_1 + 3\boldsymbol{v}_2$，代入得

$$\begin{aligned} \boldsymbol{x} &= c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_3(5\boldsymbol{v}_1 + 3\boldsymbol{v}_2) \\ &= (c_1 + 5c_3)\boldsymbol{v}_1 + (c_2 + 3c_3)\boldsymbol{v}_2 \end{aligned}$$

于是 $\boldsymbol{x}$ 在 $\mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ 中，从而 H 中每个向量属于 $\mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$，于是 H 与 $\mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ 相同，又由于 $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ 显然是线性无关的，所以 $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ 是 H 的一个基. ■

下一个定理推广了例 7.

定理 5 （生成集定理）

令 $S = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p\}$ 是 V 中的向量集，$H = \mathrm{Span}\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p\}$.

a. 若 S 中某一个向量，比如说 $\boldsymbol{v}_k$，是 S 中其余向量的线性组合，则 S 中去掉 $\boldsymbol{v}_k$ 后形成的集合仍然可以生成 H.

b. 若 $H \neq \{\mathbf{0}\}$，则 S 的某一子集是 H 的一个基.

证

a. 若必要的话，可以重排 S 中向量的顺序，这样可以假设 $\boldsymbol{v}_p$ 是 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_{p-1}$ 的线性组合，即

$$\boldsymbol{v}_p = a_1\boldsymbol{v}_1 + \cdots + a_{p-1}\boldsymbol{v}_{p-1} \tag{3}$$

任给 H 中向量 $\boldsymbol{x}$，取适当的数 $c_1,\cdots,c_p$，$\boldsymbol{x}$ 可写成

$$\boldsymbol{x} = c_1\boldsymbol{v}_1 + \cdots + c_{p-1}\boldsymbol{v}_{p-1} + c_p\boldsymbol{v}_p \tag{4}$$

将（3）中 $\boldsymbol{v}_p$ 的表达式代入（4）式，易见 $\boldsymbol{x}$ 是 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_{p-1}$ 的线性组合. 由于 $\boldsymbol{x}$ 是 H 中任一元素，所以 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_{p-1}\}$ 生成 H .

b. 若原来的生成集 S 是线性无关的，则它已经是 H 的一个基. 若不然，S 中某一个向量可表示成其余向量的线性组合，则由（a）该向量可去掉. 这样生成集中只要还有两个或更多的向量，我们就可以重复上述过程，直到这个生成集是线性无关，从而是 H 的一个基. 假如生成集最终被缩减到一个向量，则该向量是非零向量（从而是线性无关的），因为 $H \neq \{\mathbf{0}\}$. ■

Nul*A* 和 Col*A* 的基

我们已经知道如何求能够生成一个矩阵 A 的零空间的向量了，4.2 节中的讨论指出我们的方法总可以产生一个线性无关集，从而由该方法可以得到 Nul A 的一个基.

下面两个例子给出对列空间求基的简单算法.

例 8 求 Col B 的一个基，这里

$$B = [\boldsymbol{b}_1 \;\; \boldsymbol{b}_2 \;\; \cdots \;\; \boldsymbol{b}_5] = \begin{bmatrix} 1 & 4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

解 B 的每个非主元列是主元列的线性组合，事实上，$\boldsymbol{b}_2 = 4\boldsymbol{b}_1, \boldsymbol{b}_4 = 2\boldsymbol{b}_1 - \boldsymbol{b}_3$. 由生成集定理，可以去掉 $\boldsymbol{b}_2$ 和 $\boldsymbol{b}_4$，$\{\boldsymbol{b}_1,\boldsymbol{b}_3,\boldsymbol{b}_5\}$ 仍可以生成 Col B，令

$$S = \{\boldsymbol{b}_1,\boldsymbol{b}_3,\boldsymbol{b}_5\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

因为 $\boldsymbol{b}_1 \neq 0$，同时 S 中的向量都不是其前面向量的线性组合，所以 S 是线性无关的（定理 4），从而 S 是 Col B 的一个基. ■

一个矩阵 A 若不是简化阶梯形将如何？回顾 A 的列中任何线性相关关系都可以用 $A\boldsymbol{x} = \mathbf{0}$ 的形式刻画，这里 $\boldsymbol{x}$ 是一个加权的列.（若某些列没有被包括在一特殊的相关关系中，则它们的权为零.）当 A 被行化简成矩阵 B 时，B 的列通常与 A 的列完全不同，然而，方程 $A\boldsymbol{x} = \mathbf{0}$ 与 $B\boldsymbol{x} = \mathbf{0}$ 有完全相同的解集. 若 $A = [\boldsymbol{a}_1 \cdots \boldsymbol{a}_n]$ 和 $B = [\boldsymbol{b}_1 \cdots \boldsymbol{b}_n]$，则向量方程 $x_1a_1 + \cdots + x_na_n = 0$ 和 $x_1b_1 + \cdots + x_nb_n = 0$ 也有相同的解集. 即 A 的列与 B 的列具有完全相同的线性相关关系.

例 9 可以证明矩阵

$$A=[\mathbf{a}_1\ \ \mathbf{a}_2\ \ \cdots\ \ \mathbf{a}_5]=\begin{bmatrix}1&4&0&2&-1\\3&12&1&5&5\\2&8&1&3&2\\5&20&2&8&8\end{bmatrix}$$

行等价于例 8 中的矩阵 B，求 Col A 的一个基.

解 在例 8 中，易见

$$\mathbf{b}_2=4\mathbf{b}_1\ ,\ \mathbf{b}_4=2\mathbf{b}_1-\mathbf{b}_3$$

所以可以得到

$$\mathbf{a}_2=4\mathbf{a}_1\ ,\quad \mathbf{a}_4=2\mathbf{a}_1-\mathbf{a}_3$$

经检验，这是对的，从而在挑选 Col A 的最小生成集时，可以去掉 $\mathbf{a}_2$ 和 $\mathbf{a}_4$. 事实上，$\{\mathbf{a}_1,\mathbf{a}_3,\mathbf{a}_5\}$ 一定是线性无关的，这是因为 $\mathbf{a}_1,\mathbf{a}_3,\mathbf{a}_5$ 之间的任何线性相关关系都蕴涵 $\mathbf{b}_1,\mathbf{b}_3,\mathbf{b}_5$ 之间的一个线性相关关系，但我们已知 $\{\mathbf{b}_1,\mathbf{b}_3,\mathbf{b}_5\}$ 是一个线性无关集，于是 $\{\mathbf{a}_1,\mathbf{a}_3,\mathbf{a}_5\}$ 是 Col A 的一个基，此基中我们选取的列是 A 的主元列. ■

例 8 和例 9 说明了下列有用的事实.

定理 6 *矩阵 A 的主元列构成 Col A 的一个基.*

证 一般的证明用到上面讨论的论证，令 B 是 A 的简化阶梯形，由于 B 的主元列中的任一个向量都不是其前面主元列的线性组合，故 B 中的主元列是线性无关的. 又由 A 行等价于 B，A 中列的任何线性相关关系对应 B 中列的线性相关关系，所以 A 中的主元列也是线性无关的. 同理，A 中每个非主元列是 A 中主元列的线性组合，由生成集定理，A 中非主元列可以从 Col A 的生成集中去掉，剩下的 A 的主元列是 Col A 的一个基. ■

警告 *对 Col A 的基，要慎重使用 A 本身的主元列，阶梯形 B 的主元列通常不在 A 的列空间中，比如，例 8 中 B 的列最后一个数字均为零，所以它们不能生成例 9 中 A 的列空间.*

关于基的两点观察

使用生成集定理时，从生成集中删除向量，当集合变成线性无关时必须停止，因为如果再多删一个向量，该向量将不是剩下向量的线性组合，从而这个较小的集合将不再生成 V，所以一个基是一个尽可能小的生成集.

基还是尽可能大的线性无关集，若 S 是 V 的一个基，在 S 中再添加进一个新的向量，比如是从 V 中取的一个 $\mathbf{w}$，则新的集合不再是线性无关了，这是因为 S 生成 V，因此 $\mathbf{w}$ 是 S 中元素的线性组合.

例 10 下列 $\mathbb{R}^3$ 中的三个集合说明一个线性无关集如何被扩充为一个基，同时进一步的扩充如何破坏这个集合的线性无关性. 再者，一个生成集可以收缩成一个基，但进一步的收缩就破坏了生成性.

$$\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix}\right\} \quad \left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix},\begin{bmatrix}4\\5\\6\end{bmatrix}\right\} \quad \left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix},\begin{bmatrix}4\\5\\6\end{bmatrix},\begin{bmatrix}7\\8\\9\end{bmatrix}\right\}$$

线性无关，但不能生成$\mathbb{R}^3$　　$\mathbb{R}^3$的一个基　　生成$\mathbb{R}^3$，但线性相关

■

练习题

1. 令 $\boldsymbol{v}_1=\begin{bmatrix}1\\-2\\3\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}-2\\7\\-9\end{bmatrix}$，判断 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是否是 $\mathbb{R}^3$ 的一个基，$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是否是 $\mathbb{R}^2$ 的一个基.

2. 令 $\boldsymbol{v}_1=\begin{bmatrix}1\\-3\\4\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}6\\2\\-1\end{bmatrix}$，$\boldsymbol{v}_3=\begin{bmatrix}2\\-2\\3\end{bmatrix}$，$\boldsymbol{v}_4=\begin{bmatrix}-4\\-8\\9\end{bmatrix}$，求由 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4\}$ 生成的子空间 W 的一个基.

3. 令 $\boldsymbol{v}_1=\begin{bmatrix}1\\0\\0\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}0\\1\\0\end{bmatrix}$，$H=\left\{\begin{bmatrix}s\\s\\0\end{bmatrix}:s\in\mathbb{R}\right\}$，因为 $\begin{bmatrix}s\\s\\0\end{bmatrix}=s\begin{bmatrix}1\\0\\0\end{bmatrix}+s\begin{bmatrix}0\\1\\0\end{bmatrix}$ 知 H 中每一个向量都是 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的线性组合，问 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是 H 的一个基吗？

习题 4.3

判断习题 1~8 中哪一个集合是 $\mathbb{R}^2$ 或 $\mathbb{R}^3$ 的基，在不是基的集合中，判断哪一个是线性无关的，哪一个能生成 $\mathbb{R}^2$ 或 $\mathbb{R}^3$，证明你的答案.

1. $\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}1\\1\\1\end{bmatrix}$

2. $\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix}$

3. $\begin{bmatrix}1\\0\\-3\end{bmatrix},\begin{bmatrix}3\\1\\-4\end{bmatrix},\begin{bmatrix}-2\\-1\\1\end{bmatrix}$

4. $\begin{bmatrix}2\\-1\\1\end{bmatrix},\begin{bmatrix}2\\-3\\2\end{bmatrix},\begin{bmatrix}-8\\5\\4\end{bmatrix}$

5. $\begin{bmatrix}3\\-3\\0\end{bmatrix},\begin{bmatrix}-3\\7\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}0\\-3\\5\end{bmatrix}$

6. $\begin{bmatrix}1\\2\\-4\end{bmatrix},\begin{bmatrix}-4\\3\\6\end{bmatrix}$

7. $\begin{bmatrix}-2\\3\\0\end{bmatrix},\begin{bmatrix}6\\-1\\5\end{bmatrix}$

8. $\begin{bmatrix}1\\-2\\3\end{bmatrix},\begin{bmatrix}0\\3\\-1\end{bmatrix},\begin{bmatrix}2\\-1\\5\end{bmatrix},\begin{bmatrix}0\\0\\-1\end{bmatrix}$

对习题 9~10 中的矩阵，求其零空间的基，参考 4.2 节中例 3 下面的讨论.

9. $\begin{bmatrix}1&0&-2&-2\\0&1&1&4\\3&-1&-7&3\end{bmatrix}$

10. $\begin{bmatrix}1&1&-2&1&5\\0&1&0&-1&-2\\0&0&-8&0&16\end{bmatrix}$

11. 求 $\mathbb{R}^3$ 中平面 $x-3y+2z=0$ 中向量的集合的一个基.（提示：将该方程视为一个齐次线性方程组.）

12. 求 $\mathbb{R}^2$ 中直线 $y=-3x$ 上向量的集合的一个基.

在习题 13~14 中，假设 A 行等价于 B，求 $\mathrm{Nul}\,A$ 和 $\mathrm{Col}\,A$ 的基.

13. $A=\begin{bmatrix}-2&4&-2&-4\\2&-6&-3&1\\-3&8&2&-3\end{bmatrix}$，$B=\begin{bmatrix}1&0&6&5\\0&2&5&3\\0&0&0&0\end{bmatrix}$

14. $A=\begin{bmatrix}1&2&3&-4&8\\1&2&0&2&8\\2&4&-3&10&9\\3&6&0&6&9\end{bmatrix}$，$B=\begin{bmatrix}1&2&0&2&5\\0&0&3&-6&3\\0&0&0&0&-7\\0&0&0&0&0\end{bmatrix}$

在习题 15~18 中，求由给出向量 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_5$ 生成的空间的一个基.

15. $\begin{bmatrix}1\\0\\-2\\3\end{bmatrix},\begin{bmatrix}0\\1\\2\\3\end{bmatrix},\begin{bmatrix}-2\\-2\\-8\\0\end{bmatrix},\begin{bmatrix}2\\-1\\10\\3\end{bmatrix},\begin{bmatrix}3\\-1\\-6\\9\end{bmatrix}$

16. $\begin{bmatrix}1\\0\\0\\1\end{bmatrix}, \begin{bmatrix}-2\\0\\0\\2\end{bmatrix}, \begin{bmatrix}3\\-1\\1\\-1\end{bmatrix}, \begin{bmatrix}5\\-3\\3\\-4\end{bmatrix}, \begin{bmatrix}2\\-1\\1\\0\end{bmatrix}$

17. [M] $\begin{bmatrix}2\\0\\-4\\-6\\0\end{bmatrix}, \begin{bmatrix}4\\0\\2\\-4\\4\end{bmatrix}, \begin{bmatrix}-2\\-4\\0\\1\\-7\end{bmatrix}, \begin{bmatrix}8\\4\\8\\-3\\15\end{bmatrix}, \begin{bmatrix}-8\\4\\0\\0\\1\end{bmatrix}$

18. [M] $\begin{bmatrix}-3\\2\\6\\0\\-7\end{bmatrix}, \begin{bmatrix}3\\0\\-9\\0\\6\end{bmatrix}, \begin{bmatrix}0\\2\\-4\\0\\-1\end{bmatrix}, \begin{bmatrix}6\\-2\\-14\\0\\13\end{bmatrix}, \begin{bmatrix}-6\\3\\0\\-1\\0\end{bmatrix}$

19. 令 $\boldsymbol{v}_1=\begin{bmatrix}4\\-3\\7\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}1\\9\\-2\end{bmatrix}$，$\boldsymbol{v}_3=\begin{bmatrix}7\\11\\6\end{bmatrix}$，

$H=\mathrm{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 可以证明 $4\boldsymbol{v}_1+5\boldsymbol{v}_2-3\boldsymbol{v}_3=\boldsymbol{0}$，利用这个信息求 H 的一个基，答案不唯一.

20. 令 $\boldsymbol{v}_1=\begin{bmatrix}3\\4\\-2\\-5\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}4\\3\\2\\4\end{bmatrix}$，$\boldsymbol{v}_3=\begin{bmatrix}2\\5\\-6\\-14\end{bmatrix}$，可以证明

$2\boldsymbol{v}_1-\boldsymbol{v}_2-\boldsymbol{v}_3=\boldsymbol{0}$，利用此信息求 $H=\mathrm{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 的一个基.

在习题 21 和习题 22 中，标出每个命题的真假，给出理由.

21. a. 单独一个向量是线性相关的.
    b. 若 $H=\mathrm{Span}\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_p\}$，则 $\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_p\}$ 是 H 的一个基.
    c. 一个 $n\times n$ 可逆矩阵的列构成 $\mathbb{R}^n$ 的一个基.
    d. 基是一个尽可能大的生成集.
    e. 在某些情形，矩阵列之间的线性相关关系受某种初等行变换的影响.

22. a. 子空间 H 中一个线性无关集是 H 的一个基.
    b. 若非零向量的一个有限集合 S 生成一个向量空间 V，则 S 的某子集是 V 的一个基.
    c. 基是一个尽可能大的线性无关集.
    d. 4.2 节中描述的对 $\mathrm{Nul}\,A$ 产生一个生成集的标准方法，有时对产生 $\mathrm{Nul}\,A$ 的基不起作用.
    e. 若 B 是矩阵 A 的阶梯形，则 B 的主元列构成 $\mathrm{Col}\,A$ 的一个基.

23. 设 $\mathbb{R}^4=\mathrm{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_4\}$，解释为什么 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_4\}$ 是 $\mathbb{R}^4$ 的一个基.

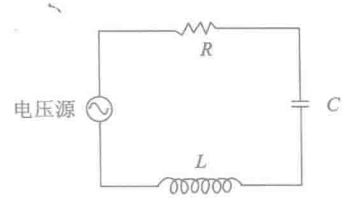
24. 令 $\mathcal{B}=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$ 是 $\mathbb{R}^n$ 中的一个线性无关集，解释为什么 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$ 是 $\mathbb{R}^n$ 的一个基.

25. 令 $\boldsymbol{v}_1=\begin{bmatrix}1\\0\\1\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}0\\1\\1\end{bmatrix}$，$\boldsymbol{v}_3=\begin{bmatrix}0\\1\\0\end{bmatrix}$，令 H 是 $\mathbb{R}^3$ 中第二和第三个数字相同的向量的集合，由于 $\begin{bmatrix}s\\t\\t\end{bmatrix}=s\begin{bmatrix}1\\0\\1\end{bmatrix}+(t-s)\begin{bmatrix}0\\1\\1\end{bmatrix}+s\begin{bmatrix}0\\1\\0\end{bmatrix}$ 对任意 s 和 t 均成立，则 H 中每个向量有由 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 线性组合而成的唯一的一个表达式，问 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 是 H 的基吗？为什么是或为什么不是？

26. 在所有实函数的向量空间中，求由 $\{\sin t,\sin 2t,\sin t\cos t\}$ 生成的子空间的一个基.

27. 令 V 是刻画物体-弹簧系统振动的函数的向量空间，（参考 4.1 节中习题 19），求 V 的一个基.

28. （*RLC* 电路）下图中的电路由一个电阻器[R 欧姆]、一个感应器[L 亨]、一个电容器[C 法拉]和一个初始电压源组成. 令 $b=R/(2L)$，同时假设 R，L，C 已经选好使得 b 也等于 $1/\sqrt{LC}$.（这是可以做到的，比如，在一个伏特计中使用该电路时.）令 $v(t)$ 表示时间 t 时的电压（伏特），它由电容器的两端测得. 可以证明 v 在将 $v(t)$ 映到 $Lv''(t)+Rv'(t)+(1/C)v(t)$ 的线性变换的零空间 H 中，H 由所有形如 $v(t)=\mathrm{e}^{-bt}(c_1+c_2t)$ 的函数构成，求 H 的一个基.



习题 29~30 表明 $\mathbb{R}^n$ 中每一个基一定恰好由 n 个向量构成.

29. 令 $S=\{v_1,\cdots,v_k\}$ 是 $\mathbb{R}^n$ 中 k 个向量的集合，$k<n$，利用 1.4 节中的一个定理解释为什么 S 不能是 $\mathbb{R}^n$ 的一个基.
30. 令 $S=\{v_1,\cdots,v_k\}$ 是 $\mathbb{R}^n$ 中 k 个向量的集合，$k>n$，利用第 1 章中的一个定理解释为什么 S 不能是 $\mathbb{R}^n$ 的一个基.

习题 31 和习题 32 表明线性无关和线性变换之间的一个重要联系，同时提供了使用线性相关定义的练习，令 V 和 W 是向量空间，令 $T:V\to W$ 是一个线性变换，且令 $\{v_1,\cdots,v_p\}$ 是 V 的一个子集.

31. 证明：若 $\{v_1,\cdots,v_p\}$ 在 V 中是线性相关的，则其像集 $\{T(v_1),\cdots,T(v_p)\}$ 在 W 中也是线性相关的. 这个事实表明如果一个线性变换将集 $\{v_1,\cdots,v_p\}$ 映射成线性无关集 $\{T(v_1),\cdots,T(v_p)\}$，则原来的集合也是线性无关的.（因为它不能是线性相关的.）
32. 假设 T 是一个一一的变换，使得方程 $T(u)=T(v)$ 总是蕴涵 $u=v$，证明：如果像集 $\{T(v_1),\cdots,T(v_p)\}$ 是线性相关的，则 $\{v_1,\cdots,v_p\}$ 也线性相关. 这个事实表明一个一对一的线性变换将线性无关集映射到一个线性无关集.（因为在此种情形下，像集合不能是线性相关的.）
33. 考虑多项式 $p_1(t)=1+t^2$，$p_2(t)=1-t^2$，$\{p_1,p_2\}$ 是 $\mathbb{P}_3$ 中的线性无关集吗？为什么？
34. 考虑多项式 $p_1(t)=1+t$，$p_2(t)=1-t$，$p_3(t)=2$（对所有的 t），检查 p_1，p_2，p_3之间的线性相关关系，然后求 $\mathrm{Span}\{p_1,p_2,p_3\}$ 的一个基.
35. 设 V 是一个包含向量集 $\{u_1,u_2,u_3,u_4\}$ 的向量空间，说明如何构造 V 中的一个向量集 $\{v_1,v_2,v_3,v_4\}$ 使得 $\{v_1,v_3\}$ 是 $\mathrm{Span}\{v_1,v_2,v_3,v_4\}$ 的一组基.
36. [M]设 $H=\mathrm{Span}\{u_1,u_2,u_3\}$，$K=\mathrm{Span}\{v_1,v_2,v_3\}$，其中

$$u_1=\begin{bmatrix}1\\2\\0\\-1\end{bmatrix},\ u_2=\begin{bmatrix}0\\2\\-1\\1\end{bmatrix},\ u_3=\begin{bmatrix}3\\4\\1\\-4\end{bmatrix},$$
$$v_1=\begin{bmatrix}-2\\-2\\-1\\3\end{bmatrix},\ v_2=\begin{bmatrix}2\\3\\2\\-6\end{bmatrix},\ v_3=\begin{bmatrix}-1\\4\\6\\-2\end{bmatrix}$$

求 H,K 和 $H+K$ 的基.（参见 4.1 节的习题 33 和习题 34.）
37. [M]证明：$\{t,\sin t,\cos 2t,\sin t\cos t\}$ 是定义在 $\mathbb{R}$ 上的函数集合的一个线性无关集，开始先假设
$$c_1\cdot t+c_2\cdot\sin t+c_3\cdot\cos 2t+c_4\cdot\sin t\cos t=0 \quad (5)$$
方程（5）对所有实数 t 一定均成立，选几个特殊的 t 值（比如 $t=0,0.1,0.2$）直到得到足够的方程构成的方程组以便确定所有 c_j 一定为零.
38. [M]证明：$\{1,\cos t,\cos^2 t,\cdots,\cos^6 t\}$ 是定义在 $\mathbb{R}$ 上的函数集合的一个线性无关集，利用 37 题中的方法.（此结果在 4.5 节中的习题 34 中用到.）

练习题答案

1. 令 $A=[v_1\ \ v_2]$，行变换表明

$$A=\begin{bmatrix}1&-2\\-2&7\\3&-9\end{bmatrix}\sim\begin{bmatrix}1&-2\\0&3\\0&0\end{bmatrix}$$

A 中的行并不是每一个都含有一个主元位置，由 1.4 节中的定理 4，A 的列不能生成 $\mathbb{R}^3$，从而 $\{v_1,v_2\}$ 不是 $\mathbb{R}^3$ 的一个基. 又由于 v_1 和 v_2 不在 $\mathbb{R}^2$ 中，它们不可能是 $\mathbb{R}^2$ 的一个基. 然而，由于 v_1 和 v_2 显然是线性无关的，它们是 $\mathbb{R}^3$ 的子空间 $\mathrm{Span}\{v_1,v_2\}$ 的一个基.

2. 构造矩阵 A，它的列空间由 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4\}$ 生成，然后行化简 A 从而找到它的主元列.

$$A=\begin{bmatrix}1 & 6 & 2 & -4\\ -3 & 2 & -2 & -8\\ 4 & -1 & 3 & 9\end{bmatrix}\sim\begin{bmatrix}1 & 6 & 2 & -4\\ 0 & 20 & 4 & -20\\ 0 & -25 & -5 & 25\end{bmatrix}\sim\begin{bmatrix}1 & 6 & 2 & -4\\ 0 & 5 & 1 & -5\\ 0 & 0 & 0 & 0\end{bmatrix}$$

A 的前两列是主元列，从而构成 $\text{Col}\,A=W$ 的一个基，从而 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是 W 的一个基.

注意：为了确定主元列，不需要将 A 化成简化阶梯形.

3. $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 均不在 H 中，所以 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 不能是 H 的一个基. 事实上，$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是所有形如 $(c_1,c_2,0)$ 的向量构成的平面的一个基，而 H 仅仅是一条直线.

4.4 坐标系

对一个向量空间 V，明确指定一个基 $\mathcal{B}$ 的一个重要原因是在 V 上强加一个“坐标系”. 本节我们将证明如果 $\mathcal{B}$ 包含 n 个向量，则坐标系将使 V 像 $\mathbb{R}^n$ 一样便于操作. 若 V 就是 $\mathbb{R}^n$ 本身，则 $\mathcal{B}$ 将确定一个坐标系，它给 V 以一个新的“视野”.

坐标系的存在依靠下列基本结果.

定理 7 （唯一表示定理）

令 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ 是向量空间 V 的一个基，则对 V 中每个向量 $\boldsymbol{x}$，存在唯一的一组数 $c_1,\cdots,c_n$ 使得

$$\boldsymbol{x}=c_1\boldsymbol{b}_1+\cdots+c_n\boldsymbol{b}_n \tag{1}$$

证 由于 $\mathcal{B}$ 生成 V，则存在一组数 $c_1,\cdots,c_n$ 使得（1）式成立，假设 $\boldsymbol{x}$ 还有其他表示：

$$\boldsymbol{x}=d_1\boldsymbol{b}_1+\cdots+d_n\boldsymbol{b}_n$$

$d_1,\cdots,d_n$ 为数，则二式相减有

$$\boldsymbol{0}=\boldsymbol{x}-\boldsymbol{x}=(c_1-d_1)\boldsymbol{b}_1+\cdots+(c_n-d_n)\boldsymbol{b}_n \tag{2}$$

由于 $\mathcal{B}$ 是线性无关的，（2）中的权一定为零，即 $c_j=d_j$ 对 $1\leqslant j\leqslant n$ 均成立. ■

定义 假设集合 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ 是 V 的一个基，$\boldsymbol{x}$ 在 V 中，$\boldsymbol{x}$ 相对于基 $\mathcal{B}$ 的坐标（或 $\boldsymbol{x}$ 的 $\mathcal{B}$-坐标）是使得 $\boldsymbol{x}=c_1\boldsymbol{b}_1+\cdots+c_n\boldsymbol{b}_n$ 的权 $c_1,\cdots,c_n$.

若 $c_1,\cdots,c_n$ 是 $\boldsymbol{x}$ 的 $\mathcal{B}$-坐标，则 $\mathbb{R}^n$ 中的向量

$$[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}$$

是 $\boldsymbol{x}$（相对于 $\mathcal{B}$）的坐标向量，或 $\boldsymbol{x}$ 的 $\mathcal{B}$-坐标向量，映射 $\boldsymbol{x}\mapsto[\boldsymbol{x}]_{\mathcal{B}}$ 称为（由 $\mathcal{B}$ 确定的）坐标映射.[1]

---

[1] 在坐标映射的概念中假定基 $\mathcal{B}$ 是一组有编号的向量集，其向量以一个预先指定的次序排列. 这个性质使得 $[\boldsymbol{x}]_{\mathcal{B}}$ 的定义没有二义性.

例 1　考虑 $\mathbb{R}^2$ 的一个基 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$，这里 $\boldsymbol{b}_1=\begin{bmatrix}1\\0\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}1\\2\end{bmatrix}$，假设 $\mathbb{R}^2$ 中一向量 $\boldsymbol{x}$ 具有坐标向量 $[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}-2\\3\end{bmatrix}$，求 $\boldsymbol{x}$.

解　$\boldsymbol{x}$ 的 $\mathcal{B}$ – 坐标揭示如何由 $\mathcal{B}$ 中的向量求 $\boldsymbol{x}$，即 $\boldsymbol{x}=(-2)\boldsymbol{b}_1+3\boldsymbol{b}_2=(-2)\begin{bmatrix}1\\0\end{bmatrix}+3\begin{bmatrix}1\\2\end{bmatrix}=\begin{bmatrix}1\\6\end{bmatrix}$. ■

例 2　向量 $\boldsymbol{x}=\begin{bmatrix}1\\6\end{bmatrix}$ 中的数字是 $\boldsymbol{x}$ 相对于标准基 $\mathcal{E}=\{\boldsymbol{e}_1,\boldsymbol{e}_2\}$ 的坐标，由于

$$\begin{bmatrix}1\\6\end{bmatrix}=1\cdot\begin{bmatrix}1\\0\end{bmatrix}+6\cdot\begin{bmatrix}0\\1\end{bmatrix}=1\cdot\boldsymbol{e}_1+6\cdot\boldsymbol{e}_2$$

若 $\mathcal{E}=\{\boldsymbol{e}_1,\boldsymbol{e}_2\}$，则 $[\boldsymbol{x}]_{\mathcal{E}}=\boldsymbol{x}$. ■

坐标的几何意义

一个集合上的坐标系由此集合中点到 $\mathbb{R}^n$ 中的一一映射组成. 例如，当我们选取垂直的轴同时在每个轴上取一个相同的度量单位时，通常的图纸给出了平面上的一个坐标系. 图 4-15 展示了标准基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2\}$，例 1 中的向量 $\boldsymbol{b}_1(=\boldsymbol{e}_1)$ 和 $\boldsymbol{b}_2$ 以及向量 $\boldsymbol{x}=\begin{bmatrix}1\\6\end{bmatrix}$. 坐标 1 和 6 给出 $\boldsymbol{x}$ 相对于标准基的位置：在 $\boldsymbol{e}_1$ 方向上有 1 个单位，在 $\boldsymbol{e}_2$ 方向上有 6 个单位.

图 4-16 展示来自图 4-15 的向量 $\boldsymbol{b}_1,\boldsymbol{b}_2$ 和 $\boldsymbol{x}$.（从几何意义上看，这三个向量在这两个图中均位于一条垂线上.）然而，标准坐标的格子被去掉同时被特别适合例 1 中的坐标 $\mathcal{B}$ 的格子所取代. 坐标向量 $[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}-2\\3\end{bmatrix}$ 给出 $\boldsymbol{x}$ 在新的坐标系中的位置：在 $\boldsymbol{b}_1$ 方向上有–2 个单位，在 $\boldsymbol{b}_2$ 方向上有 3 个单位.

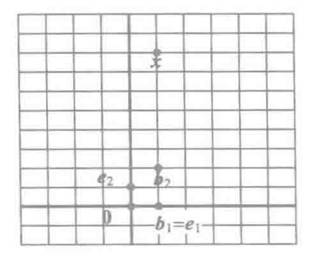


图 4-15　标准图纸

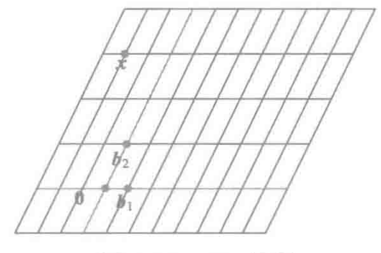
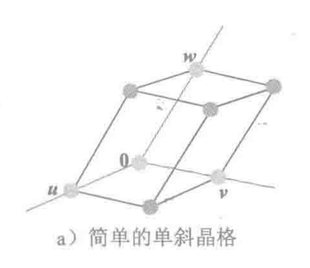


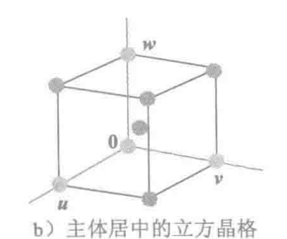
图 4-16　$\mathcal{B}$ – 图纸

例 3　在结晶学中，晶体格的刻画可由选取 $\mathbb{R}^3$ 中的一个基 $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$ 而得到帮助，这个基对应着晶体的单位方格的三个相邻的棱，一个完整的格子框架可以通过将许多个单位方格的复制品堆积在一起而构造. 有 14 种类型的单位方格，图 4-17 中展示了其中的 3 种.[㊀]

㊀ 参见 *The Science and Engineering of Materials*,4th Ed., Donald R.Askeland (Boston:Prindle, Weber & Schmidt, 2002), p.36.



a）简单的单斜晶格



b）主体居中的立方晶格

c）表面居中的正交晶格

图 4-17　单位方格的例子

相对于晶体格的基，可以给出晶体中原子的坐标. 例如$\begin{bmatrix}1/2\\1/2\\1\end{bmatrix}$标识图 4-17c 中最上面的中心原子. ■

$\mathbb{R}^n$ 中的坐标

当 $\mathbb{R}^n$ 中的一组基 $\mathcal{B}$ 固定时，容易求出任一指定的向量 $\boldsymbol{x}$ 的 $\mathcal{B}-$坐标，如下面例子所示.

例 4　令 $\boldsymbol{b}_1=\begin{bmatrix}2\\1\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}-1\\1\end{bmatrix},\boldsymbol{x}=\begin{bmatrix}4\\5\end{bmatrix},\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$，求出 $\boldsymbol{x}$ 相对于 $\mathcal{B}$ 的坐标向量 $[\boldsymbol{x}]_{\mathcal{B}}$.

解　$\boldsymbol{x}$ 的 $\mathcal{B}-$坐标 c_1,c_2 满足

$$c_1\underset{\boldsymbol{b}_1}{\begin{bmatrix}2\\1\end{bmatrix}}+c_2\underset{\boldsymbol{b}_2}{\begin{bmatrix}-1\\1\end{bmatrix}}=\underset{\boldsymbol{x}}{\begin{bmatrix}4\\5\end{bmatrix}}$$

或

$$\begin{bmatrix}2&-1\\1&1\end{bmatrix}\begin{bmatrix}c_1\\c_2\end{bmatrix}=\begin{bmatrix}4\\5\end{bmatrix}\tag{3}$$

$$\boldsymbol{b}_1\quad\boldsymbol{b}_2\qquad\qquad\boldsymbol{x}$$

这个方程可以通过在一增广矩阵上做行变换或利用左边矩阵的逆解出. 不论哪种解法，其解均为 $c_1=3,c_2=2$，从而 $\boldsymbol{x}=3\boldsymbol{b}_1+2\boldsymbol{b}_2$，同时有

$$[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}c_1\\c_2\end{bmatrix}=\begin{bmatrix}3\\2\end{bmatrix}$$

见图 4-18.

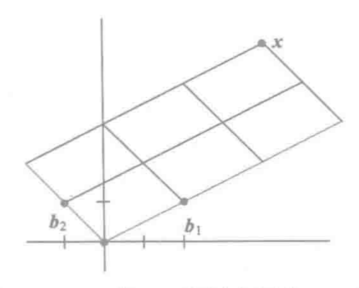


图 4-18　$\boldsymbol{x}$ 的 $\mathcal{B}-$坐标向量为（3,2） ■

（3）式中的矩阵将向量 $\boldsymbol{x}$ 的 $\mathcal{B}$ – 坐标变为 $\boldsymbol{x}$ 的标准坐标. 对 $\mathbb{R}^n$ 中的一个基 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$，可以施行类似的坐标变换. 令

$$P_{\mathcal{B}}=[\boldsymbol{b}_1\ \ \boldsymbol{b}_2\ \ \cdots\ \ \boldsymbol{b}_n]$$

则向量方程

$$\boldsymbol{x}=c_1\boldsymbol{b}_1+c_2\boldsymbol{b}_2+\cdots+c_n\boldsymbol{b}_n$$

等价于

$$\boxed{\boldsymbol{x}=P_{\mathcal{B}}[\boldsymbol{x}]_{\mathcal{B}}} \tag{4}$$

我们称 $P_{\mathcal{B}}$ 为从 $\mathcal{B}$ 到 $\mathbb{R}^n$ 中标准基的**坐标变换矩阵**. 通过左乘 $P_{\mathcal{B}}$ 将坐标向量 $[\boldsymbol{x}]_{\mathcal{B}}$ 变换到 $\boldsymbol{x}$. 坐标变换方程（4）是重要的，它将在第 5 章和第 7 章的多个地方用到.

由于 $P_{\mathcal{B}}$ 的列构成 $\mathbb{R}^n$ 的一个基，$P_{\mathcal{B}}$ 是可逆的（由可逆矩阵定理）. 通过左乘 $P_{\mathcal{B}}^{-1}$ 又将 $\boldsymbol{x}$ 变回 $\mathcal{B}$ – 坐标向量：

$$P_{\mathcal{B}}^{-1}\boldsymbol{x}=[\boldsymbol{x}]_{\mathcal{B}}$$

这里由 $P_{\mathcal{B}}^{-1}$ 产生的映射 $\boldsymbol{x}\mapsto[\boldsymbol{x}]_{\mathcal{B}}$ 是前面提到的坐标映射. 因为 $P_{\mathcal{B}}^{-1}$ 是一可逆矩阵，由可逆矩阵定理（也可参见 1.9 节定理 12），此坐标映射是一个由 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 上的一一线性变换. 我们将会看到，坐标映射的这个性质对具有一个基的一般的向量空间也成立.

坐标映射

对向量空间 V 选定一个基 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$，它引出 V 中一个坐标系. 坐标映射 $\boldsymbol{x}\mapsto[\boldsymbol{x}]_{\mathcal{B}}$ 将可能不熟悉的空间 V 与熟悉的空间 $\mathbb{R}^n$ 联系了起来，见图 4-19. V 中的点现在可以由它们的新“名字”来确定.

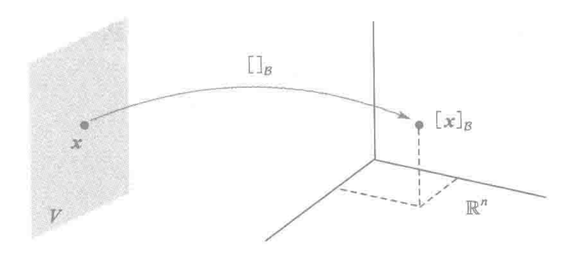


图 4-19　由 V 映上到 $\mathbb{R}^n$ 的坐标映射

定理 8　令 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ 是向量空间 V 的一个基，则坐标映射 $\boldsymbol{x}\mapsto[\boldsymbol{x}]_{\mathcal{B}}$ 是一个由 V 映上到 $\mathbb{R}^n$ 的一对一的线性变换.

证　取 V 中两个典型的向量，比如

$$\boldsymbol{u}=c_1\boldsymbol{b}_1+\cdots+c_n\boldsymbol{b}_n$$
$$\boldsymbol{w}=d_1\boldsymbol{b}_1+\cdots+d_n\boldsymbol{b}_n$$

利用向量运算，

$$\boldsymbol{u}+\boldsymbol{w}=(c_1+d_1)\boldsymbol{b}_1+\cdots+(c_n+d_n)\boldsymbol{b}_n$$

于是

$$[\boldsymbol{u}+\boldsymbol{w}]_{\mathcal{B}}=\begin{bmatrix}c_1+d_1\\ \vdots\\ c_n+d_n\end{bmatrix}=\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}+\begin{bmatrix}d_1\\ \vdots\\ d_n\end{bmatrix}=[\boldsymbol{u}]_{\mathcal{B}}+[\boldsymbol{w}]_{\mathcal{B}}$$

从而坐标映射保持加法封闭. 若 r 是任一数，则

$$r\boldsymbol{u}=r(c_1\boldsymbol{b}_1+\cdots+c_n\boldsymbol{b}_n)=(rc_1)\boldsymbol{b}_1+\cdots+(rc_n)\boldsymbol{b}_n$$

于是

$$[r\boldsymbol{u}]_{\mathcal{B}}=\begin{bmatrix}rc_1\\ \vdots\\ rc_n\end{bmatrix}=r\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}=r[\boldsymbol{u}]_{\mathcal{B}}$$

从而坐标映射也保持标量乘法封闭. 于是坐标映射是一个线性变换. 坐标映射是一对一的以及它将 V 映上到 $\mathbb{R}^n$ 的证明分别留作习题 23 和 24. ■

正如 1.8 节那样，坐标映射的线性可推广到线性组合. 若 $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p$ 在 V 中，$c_1,\cdots,c_p$ 是数，则

$$[c_1\boldsymbol{u}_1+\cdots+c_p\boldsymbol{u}_p]_{\mathcal{B}}=c_1[\boldsymbol{u}_1]_{\mathcal{B}}+\cdots+c_p[\boldsymbol{u}_p]_{\mathcal{B}} \tag{5}$$

换句话说，（5）说明 $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p$ 的一个线性组合的 $\mathcal{B}$－坐标向量等于它们坐标向量的相同的线性组合.

定理 8 中的坐标映射是一个由 V 到 $\mathbb{R}^n$ 上同构的重要例子. 一般而言，从一个向量空间 V 映上到另一个向量空间 W 的一一线性变换称为从 V 和 W 上的一个**同构**（isomorphism）.（在希腊语中 iso 表示相同，morph 表示形状或结构.）V 和 W 的记号和术语可能不同，但这两个空间作为向量空间则不加以区分. *每一个在 V 中的向量空间的计算可以完全相同地出现在 W 中，反之亦然.* 见习题 25 和习题 26.

例 5 令 $\mathcal{B}$ 是多项式空间 $\mathbb{P}_3$ 的标准基，即 $\mathcal{B}=\{1,t,t^2,t^3\}$，$\mathbb{P}_3$ 中的一个典型元素 $\boldsymbol{p}$ 具有形式：

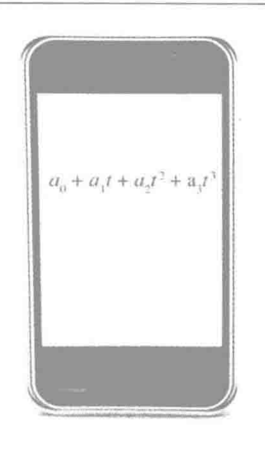
$$\boldsymbol{p}(t)=a_0+a_1t+a_2t^2+a_3t^3$$

因 $\boldsymbol{p}$ 已经给出了标准基向量的一个线性组合，我们断定

$$[\boldsymbol{p}]_{\mathcal{B}}=\begin{bmatrix}a_0\\ a_1\\ a_2\\ a_3\end{bmatrix}$$

于是坐标映射 $\boldsymbol{p}\mapsto[\boldsymbol{p}]_{\mathcal{B}}$ 是一个 $\mathbb{P}_3$ 到 $\mathbb{R}^4$ 上的同构. $\mathbb{P}_3$ 中所有向量空间运算都对应着 $\mathbb{R}^4$ 中的运算. ■

如果我们考虑将 $\mathbb{P}_3$ 和 $\mathbb{R}^4$ 分别展现在两个计算机的显示器上，两个显示器由坐标变换相联系，则一个显示器中 $\mathbb{P}_3$ 的每一个向量空间的运算被正确地复制到另一个显示器中 $\mathbb{R}^4$ 的一个对应的向量运算. $\mathbb{P}_3$ 显示器上的向量看起来与 $\mathbb{R}^4$ 显示器上的向量不同，但它们作为向量的“作用”是完全相同的，见图 4-20.



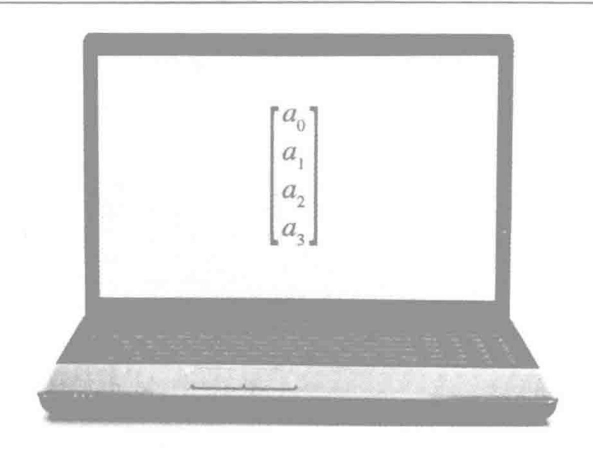


图 4-20 空间 $\mathbb{P}_3$ 与 $\mathbb{R}^4$ 同构

例 6 利用坐标向量证明在 $\mathbb{P}_2$ 中多项式 $1+2t^2, 4+t+5t^2, 3+2t$ 是线性相关的.

解 由例 5 中的坐标映射可分别产生出坐标向量（1,0,2），（4,1,5），（3,2,0）．将这些向量写成一个矩阵 A 的列，可以通过行化简 $A\boldsymbol{x}=\boldsymbol{0}$ 的增广矩阵来断定它们的线性相关性：

$$\begin{bmatrix} 1 & 4 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 2 & 5 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

A 的列是线性相关的，所以对应的多项式也是线性相关的. 事实上，容易检查 A 的第 3 列是 2 倍的第 2 列减去 5 倍的第 1 列. 多项式的对应关系是

$$3+2t = 2(4+t+5t^2) - 5(1+2t^2)$$

■

最后一个例子是关于 $\mathbb{R}^3$ 中一个与 $\mathbb{R}^2$ 同构的平面.

例 7 令 $\boldsymbol{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}, \mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2\}$，则 $\mathcal{B}$ 是 $H = \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ 的一个基. 判定 $\boldsymbol{x}$ 是否在 H 中，若在，求 $\boldsymbol{x}$ 相对于 $\mathcal{B}$ 的坐标向量.

解 若 $\boldsymbol{x}$ 在 H 中，则下列向量方程是相容的：

$$c_1 \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$$

数 c_1, c_2 若存在的话，它们就是 $\boldsymbol{x}$ 的 $\mathcal{B}$-坐标. 利用行变换，得到

$$\begin{bmatrix} 3 & -1 & 3 \\ 6 & 0 & 12 \\ 2 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

于是 $c_1 = 2, c_2 = 3, [\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$，由 $\mathcal{B}$ 确定的 H 上的坐标系如图 4-21 所示.

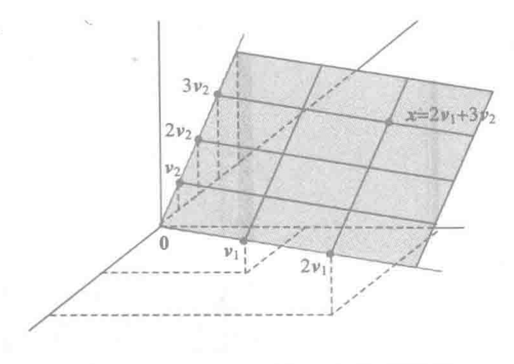


图 4-21 $\mathbb{R}^3$ 中平面 H 上的坐标系 ■

若 H 选一个不同的基，相应的坐标系也使 H 与 $\mathbb{R}^2$ 同构吗？这一定是正确的，下节中我们将证明这一结论.

练习题

1. 令 $\boldsymbol{b}_1=\begin{bmatrix}1\\0\\0\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}-3\\4\\0\end{bmatrix},\boldsymbol{b}_3=\begin{bmatrix}3\\-6\\3\end{bmatrix},\boldsymbol{x}=\begin{bmatrix}-8\\2\\3\end{bmatrix}$.
   a. 证明集合 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3\}$ 是 $\mathbb{R}^3$ 的一个基.
   b. 求由 $\mathcal{B}$ 到标准基的坐标变换矩阵.
   c. 写出 $\mathbb{R}^3$ 中 $\boldsymbol{x}$ 与 $[\boldsymbol{x}]_{\mathcal{B}}$ 相关联的方程.
   d. 对上面给出的 $\boldsymbol{x}$，求 $[\boldsymbol{x}]_{\mathcal{B}}$.
2. 集合 $\mathcal{B}=\{1+t,1+t^2,t+t^2\}$ 是 $\mathbb{P}_2$ 的一个基. 求 $\boldsymbol{p}(t)=6+3t-t^2$ 关于 $\mathcal{B}$ 的坐标向量.

习题 4.4

在习题 1~4 中，已知基 $\mathcal{B}$ 和坐标向量 $[\boldsymbol{x}]_{\mathcal{B}}$，求向量 $\boldsymbol{x}$.

1. $\mathcal{B}=\left\{\begin{bmatrix}3\\-5\end{bmatrix},\begin{bmatrix}-4\\6\end{bmatrix}\right\},[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}5\\3\end{bmatrix}$

2. $\mathcal{B}=\left\{\begin{bmatrix}3\\2\end{bmatrix},\begin{bmatrix}-4\\1\end{bmatrix}\right\},[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}-2\\5\end{bmatrix}$

3. $\mathcal{B}=\left\{\begin{bmatrix}1\\-2\\3\end{bmatrix},\begin{bmatrix}5\\0\\-2\end{bmatrix},\begin{bmatrix}4\\-3\\0\end{bmatrix}\right\},[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}1\\0\\-2\end{bmatrix}$

4. $\mathcal{B}=\left\{\begin{bmatrix}-2\\2\\0\end{bmatrix},\begin{bmatrix}3\\0\\2\end{bmatrix},\begin{bmatrix}4\\-1\\3\end{bmatrix}\right\},[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}-3\\2\\-1\end{bmatrix}$

在习题 5~8 中，求 $\boldsymbol{x}$ 关于给定基 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ 的坐标向量 $[\boldsymbol{x}]_{\mathcal{B}}$.

5. $\boldsymbol{b}_1=\begin{bmatrix}1\\-2\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}3\\-5\end{bmatrix},\boldsymbol{x}=\begin{bmatrix}-1\\1\end{bmatrix}$

6. $\boldsymbol{b}_1=\begin{bmatrix}1\\-4\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}2\\-3\end{bmatrix},\boldsymbol{x}=\begin{bmatrix}-1\\-6\end{bmatrix}$

7. $\boldsymbol{b}_1=\begin{bmatrix}1\\-1\\-3\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}-3\\4\\9\end{bmatrix},\boldsymbol{b}_3=\begin{bmatrix}2\\-2\\4\end{bmatrix},\boldsymbol{x}=\begin{bmatrix}8\\-9\\6\end{bmatrix}$

8. $\boldsymbol{b}_1=\begin{bmatrix}1\\1\\3\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}2\\0\\8\end{bmatrix},\boldsymbol{b}_3=\begin{bmatrix}1\\-1\\3\end{bmatrix},\boldsymbol{x}=\begin{bmatrix}0\\0\\-2\end{bmatrix}$

在习题 9~10 中，求由 $\mathcal{B}$ 到 $\mathbb{R}^n$ 中标准基的坐标变换矩阵.

9. $\mathcal{B}=\left\{\begin{bmatrix}1\\-3\end{bmatrix},\begin{bmatrix}2\\-5\end{bmatrix}\right\}$

10. $\mathcal{B}=\left\{\begin{bmatrix}3\\0\\6\end{bmatrix},\begin{bmatrix}2\\2\\-4\end{bmatrix},\begin{bmatrix}1\\-2\\3\end{bmatrix}\right\}$

在习题 11~12 中，对给出的 $\boldsymbol{x}$ 和 $\mathcal{B}$，利用逆矩阵求 $[\boldsymbol{x}]_{\mathcal{B}}$.

11. $\mathcal{B}=\left\{\begin{bmatrix}1\\-2\end{bmatrix},\begin{bmatrix}-3\\5\end{bmatrix}\right\},\boldsymbol{x}=\begin{bmatrix}2\\-5\end{bmatrix}$

12. $\mathcal{B}=\left\{\begin{bmatrix}1\\-1\end{bmatrix},\begin{bmatrix}2\\-1\end{bmatrix}\right\},\boldsymbol{x}=\begin{bmatrix}2\\3\end{bmatrix}$

13. 集 $\mathcal{B}=\{1+t^2,t+t^2,1+2t+t^2\}$ 是 $\mathbb{P}_2$ 的一个基，求 $\boldsymbol{p}(t)=1+4t+7t^2$ 关于 $\mathcal{B}$ 的坐标向量.

14. 集 $\mathcal{B}=\{1-t^2,t-t^2,2-t+t^2\}$ 是 $\mathbb{P}_2$ 的一个基，求 $\boldsymbol{p}(t)=1+3t-6t^2$ 关于 $\mathcal{B}$ 的坐标向量.

在习题 15 和习题 16 中，标出每个命题的真假，验证每个答案. 除非另外说明，$\mathcal{B}$ 均指向量空间 V 的一个基.

15. a. 若 $\boldsymbol{x}$ 在 V 中且 $\mathcal{B}$ 包含 n 个向量，则 $\boldsymbol{x}$ 的 $\mathcal{B}-$坐标向量在 $\mathbb{R}^n$ 中.
    b. 若 $P_{\mathcal{B}}$ 是坐标变换矩阵，则 $[\boldsymbol{x}]_{\mathcal{B}}=P_{\mathcal{B}}\boldsymbol{x},\boldsymbol{x}\in V$.
    c. 向量空间 $\mathbb{P}_3$ 和 $\mathbb{R}^3$ 是同构的.

16. a. 若 $\mathcal{B}$ 是 $\mathbb{R}^n$ 的标准基，则 $\mathbb{R}^n$ 中 $\boldsymbol{x}$ 的 $\mathcal{B}-$坐标向量是 $\boldsymbol{x}$ 本身.
    b. 对应 $[\boldsymbol{x}]_{\mathcal{B}}\mapsto\boldsymbol{x}$ 称为坐标映射.
    c. 在某种情形下，$\mathbb{R}^3$ 中的平面可以与 $\mathbb{R}^2$ 同构.

17. 向量 $\boldsymbol{v}_1=\begin{bmatrix}1\\-3\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}2\\-8\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}-3\\7\end{bmatrix}$ 生成 $\mathbb{R}^2$ 但不构成一个基，用两种不同的方法将 $\begin{bmatrix}1\\1\end{bmatrix}$ 表为 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 的线性组合.

18. 令 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ 是向量空间 V 的一个基，解释为什么 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n$ 的 $\mathcal{B}-$坐标向量是 $n\times n$ 单位矩阵的列 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n$.

19. 令 S 是向量空间 V 中的有限集，具有如下性质：V 中每个 $\boldsymbol{x}$ 均可表示为 S 中元素的唯一线性组合，证明 S 是 V 的一个基.

20. 假设 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_4\}$ 是向量空间 V 的一个线性相关生成集，证明 V 中每一个 $\boldsymbol{w}$ 都可用多于一种的方式表示成 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_4$ 的线性组合.（提示：令 $\boldsymbol{w}=k_1\boldsymbol{v}_1+\cdots+k_4\boldsymbol{v}_4$ 是 V 中一任意向量，利用 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_4\}$ 的线性相关性将 $\boldsymbol{w}$ 表示为 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_4$ 的另一个线性组合.）

21. 令 $\mathcal{B}=\left\{\begin{bmatrix}1\\-4\end{bmatrix},\begin{bmatrix}-2\\9\end{bmatrix}\right\}$，因为由 $\mathcal{B}$ 确定的坐标映射是一个从 $\mathbb{R}^2$ 映射到 $\mathbb{R}^2$ 的线性变换，此映射一定可由某 2×2 矩阵 A 来实现，求出 A.（提示：由 A 去乘可使向量 $\boldsymbol{x}$ 变换到它的坐标向量 $[\boldsymbol{x}]_{\mathcal{B}}$.）

22. 令 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ 是 $\mathbb{R}^n$ 的一个基，作出一个 $n\times n$ 矩阵 A 的描述，它使得坐标映射 $\boldsymbol{x}\mapsto[\boldsymbol{x}]_{\mathcal{B}}$ 得以实现.（见习题 21.）

习题 23~26 涉及一个向量空间 V、一个基 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ 和坐标映射 $\boldsymbol{x}\mapsto[\boldsymbol{x}]_{\mathcal{B}}$.

23. 证明：坐标映射是一一的.（提示：假设对 V 中某向量 $\boldsymbol{u}$ 和 $\boldsymbol{w}$，满足 $[\boldsymbol{u}]_{\mathcal{B}}=[\boldsymbol{w}]_{\mathcal{B}}$，证明 $\boldsymbol{u}=\boldsymbol{w}$.）

24. 证明：坐标映射是映上到 $\mathbb{R}^n$ 的，即任给 $\mathbb{R}^n$ 中向量 $\boldsymbol{y}$，具有元素 $y_1,\cdots,y_n$，存在 V 中的向量 $\boldsymbol{u}$，使得 $[\boldsymbol{u}]_{\mathcal{B}}=\boldsymbol{y}$.

25. 证明：V 中子集 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是线性无关的，当且仅当坐标向量 $\{[\boldsymbol{u}_1]_{\mathcal{B}},\cdots,[\boldsymbol{u}_p]_{\mathcal{B}}\}$ 在 $\mathbb{R}^n$ 中也是线性无关的. 提示：因为坐标映射是一一的，下列方程具有相同的解 $c_1,\cdots,c_p$.

    $c_1\boldsymbol{u}_1+\cdots+c_p\boldsymbol{u}_p=\boldsymbol{0}$　　V 中的零向量

    $[c_1\boldsymbol{u}_1+\cdots+c_p\boldsymbol{u}_p]_{\mathcal{B}}=[\boldsymbol{0}]_{\mathcal{B}}$　　$\mathbb{R}^n$ 中的零向量

26. 给定 V 中向量 $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p$ 和 $\boldsymbol{w}$，证明：$\boldsymbol{w}$ 是 $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p$ 的线性组合，当且仅当 $[\boldsymbol{w}]_{\mathcal{B}}$ 是坐标向量 $[\boldsymbol{u}_1]_{\mathcal{B}},\cdots,[\boldsymbol{u}_p]_{\mathcal{B}}$ 的线性组合.

在习题 27~30 中，利用坐标向量检验多项式集合的线性无关性.

27. $1+t^3,2+t-3t^2,-t+2t^2-t^3$

28. $1-2t^2-t^3, t+2t^3, 1+t-2t^2$

29. $(1-t)^2, t-2t^2+t^3, (1-t)^3$

30. $(2-t)^3, (3-t)^2, 1+6t-5t^2+t^3$

31. 利用坐标向量检验下面的多项式集合是否生成 $\mathbb{P}_2$，验证你的结论.

 a. $1-3t+5t^2, -3+5t-7t^2, -4+5t-6t^2, 1-t^2$

 b. $5t+t^2, 1-8t-2t^2, -3+4t+2t^2, 2-3t$

32. 设 $\boldsymbol{p}_1(t)=1+t^2$，$\boldsymbol{p}_2(t)=t-3t^2$，$\boldsymbol{p}_3(t)=1+t-3t^2$.

 a. 利用坐标向量说明这些多项式集合是 $\mathbb{P}_2$ 的一组基.

 b. 考虑 $\mathbb{P}_2$ 的一组基 $\mathcal{B}=\{\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3\}$，给定 $[\boldsymbol{q}]_\mathcal{B}=\begin{bmatrix}-1\\1\\2\end{bmatrix}$，求 $\mathbb{P}_2$ 中的 $\boldsymbol{q}$.

在习题 33~34 中，利用坐标向量检验下面的多项式集合是否构成 $\mathbb{P}_3$ 的一组基，验证你的结论.

33. [M] $3+7t, 5+t-2t^3, t-2t^2, 1+16t-6t^2+2t^3$

34. [M] $5-3t+4t^2+2t^3, 9+t+8t^2-6t^3, 6-2t+5t^2, t^3$

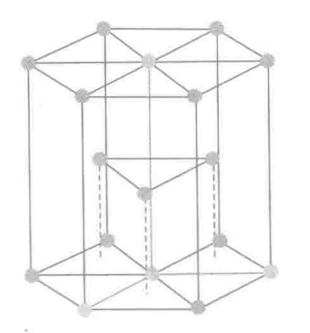
35. [M]令 $H=\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}, \mathcal{B}=\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$，证明：$\boldsymbol{x}$ 在 H 中并求 $\boldsymbol{x}$ 的 $\mathcal{B}-$坐标向量，其中

$$\boldsymbol{v}_1=\begin{bmatrix}11\\-5\\10\\7\end{bmatrix}, \boldsymbol{v}_2=\begin{bmatrix}14\\-8\\13\\10\end{bmatrix}, \boldsymbol{x}=\begin{bmatrix}19\\-13\\18\\15\end{bmatrix}$$

36. [M]令 $H=\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}, \mathcal{B}=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$，证明：$\mathcal{B}$ 是 H 的一个基并且 $\boldsymbol{x}$ 在 H 中，再求 $\boldsymbol{x}$ 的 $\mathcal{B}-$坐标向量，这里

$$\boldsymbol{v}_1=\begin{bmatrix}-6\\4\\-9\\4\end{bmatrix}, \boldsymbol{v}_2=\begin{bmatrix}8\\-3\\7\\-3\end{bmatrix}, \boldsymbol{v}_3=\begin{bmatrix}-9\\5\\-8\\3\end{bmatrix}, \boldsymbol{x}=\begin{bmatrix}4\\7\\-8\\3\end{bmatrix}$$

[M]习题 37 和习题 38 涉及钛的晶体格，它具有六边形结构，见图 4-22 左图，$\mathbb{R}^3$ 中向量 $\begin{bmatrix}2.6\\-1.5\\0\end{bmatrix}, \begin{bmatrix}0\\3\\0\end{bmatrix}, \begin{bmatrix}0\\0\\4.8\end{bmatrix}$ 对展示在右图中的单位方格构成一组基，这里的数字是以埃（1 埃 $=10^{-8}$ 厘米）为单位的，在钛合金中，一些添加的原子可能在“八面体的”和“四面体的”位置的单位方格内.（这样命名是由于这些位置的原子所形成的几何对象.）



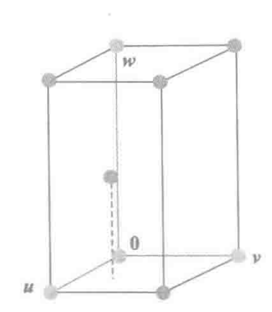


图 4-22 六边形的封装晶格和它的单位方格

37. 相对于格的基，八面体中一位置为 $\begin{bmatrix}1/2\\1/4\\1/6\end{bmatrix}$，相对于 $\mathbb{R}^3$ 的标准基，确定这个位置的坐标.

38. 四面体中一个点的位置为 $\begin{bmatrix}1/2\\1/2\\1/3\end{bmatrix}$，相对于 $\mathbb{R}^3$ 的标准基，确定这个位置的坐标.

练习题答案

1. a. 矩阵 $P_\mathcal{B}=[\boldsymbol{b}_1\ \ \boldsymbol{b}_2\ \ \boldsymbol{b}_3]$ 行等价于单位矩阵是明显的，由可逆矩阵定理，$P_\mathcal{B}$ 是可逆的，同时它的列构成 $\mathbb{R}^3$ 的一个基.

 b. 由（a）知，坐标变换矩阵为 $P_\mathcal{B}=\begin{bmatrix}1&-3&3\\0&4&-6\\0&0&3\end{bmatrix}$.

c. $\boldsymbol{x}=P_{\mathcal{B}}[\boldsymbol{x}]_{\mathcal{B}}$.

d. 为解（c），行化简增广矩阵而不计算 $P_{\mathcal{B}}^{-1}$ 可能更容易一些.

$$\underset{P_{\mathcal{B}}\qquad\qquad \boldsymbol{x}}{\begin{bmatrix}1 & -3 & 3 & -8\\0 & 4 & -6 & 2\\0 & 0 & 3 & 3\end{bmatrix}} \sim \underset{I\qquad\qquad [\boldsymbol{x}]_{\mathcal{B}}}{\begin{bmatrix}1 & 0 & 0 & -5\\0 & 1 & 0 & 2\\0 & 0 & 1 & 1\end{bmatrix}}$$

从而

$$[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}-5\\2\\1\end{bmatrix}$$

2. $\boldsymbol{p}(t)=6+3t-t^2$ 相对于 $\mathcal{B}$ 的坐标满足 $c_1(1+t)+c_2(1+t^2)+c_3(t+t^2)=6+3t-t^2$，对比 t 的同次幂项的系数有

$$\begin{aligned}c_1+c_2\qquad\quad &= 6\\c_1\qquad\quad +c_3 &= 3\\c_2+c_3 &= -1\end{aligned}$$

解之，得 $c_1=5, c_2=1, c_3=-2$，从而 $[\boldsymbol{p}]_{\mathcal{B}}=\begin{bmatrix}5\\1\\-2\end{bmatrix}$.

4.5 向量空间的维数

4.4 节定理 8.8 可推出一向量空间 V 的基 $\mathcal{B}$ 若含有 n 个向量，则 V 与 $\mathbb{R}^n$ 同构，本节我们证明数 n 是 V 的一个内在的性质（称为维数），它不依赖基的选择，维数的讨论将对基的性质有更深入的理解.

第一个定理推广关于向量空间 $\mathbb{R}^n$ 的一个著名的结果.

定理 9 若向量空间 V 具有一组基 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$，则 V 中任意包含多于 n 个向量的集合一定线性相关.

证 令 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是 V 中一个含有多于 n 个向量的集合，因为坐标向量 $[\boldsymbol{u}_1]_{\mathcal{B}},\cdots,[\boldsymbol{u}_p]_{\mathcal{B}}$ 个数（p）大于每个向量中数字的个数（n），所以 $[\boldsymbol{u}_1]_{\mathcal{B}},\cdots,[\boldsymbol{u}_p]_{\mathcal{B}}$ 线性相关. 于是存在数量 $c_1,\cdots,c_p$ 不全为 0，使得

$$c_1[\boldsymbol{u}_1]_{\mathcal{B}}+\cdots+c_p[\boldsymbol{u}_p]_{\mathcal{B}}=\begin{bmatrix}0\\\vdots\\0\end{bmatrix}\qquad \mathbb{R}^n\text{ 中的零向量}$$

因为坐标映射是线性变换

$$[c_1\boldsymbol{u}_1+\cdots+c_p\boldsymbol{u}_p]_{\mathcal{B}}=\begin{bmatrix}0\\\vdots\\0\end{bmatrix}$$

于是 $c_1\boldsymbol{u}_1+\cdots+c_p\boldsymbol{u}_p=0\cdot\boldsymbol{b}_1+\cdots+0\cdot\boldsymbol{b}_n=\boldsymbol{0}$，因为 c_i 不全为零，所以 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是线性相关的.[㊀] ■

由定理 9 可推出如果一向量空间 V 有一组 n 个向量的基，则 V 的每组基必定是由 n 个向量所组成.

定理 10 若向量空间 V 有一组基含有 n 个向量，则 V 的每一组基一定恰好含有 n 个向量.

证 令 $\mathcal{B}_1$ 是一个含 n 个向量的基，$\mathcal{B}_2$ 是 V 的任一另外的基，因为 $\mathcal{B}_1$ 是一个基，$\mathcal{B}_2$ 是线性无关的，则由定理 9，$\mathcal{B}_2$ 不能含有多于 n 个向量，同理由 $\mathcal{B}_2$ 是一个基，$\mathcal{B}_1$ 线性无关，$\mathcal{B}_2$ 至少含有 n 个向量，于是 $\mathcal{B}_2$ 恰好含有 n 个向量. ■

如果一个非零向量空间 V 由有限集 S 生成，则由生成集定理，S 的一个子集是 V 的一个基. 由此，定理 10 保证下列定义有意义.

定义 若 V 由一个有限集生成，则 V 称为**有限维的**，V 的维数写成 $\dim V$，是 V 的基中含有向量的个数，零向量空间 $\{\boldsymbol{0}\}$ 的维数定义为零. 如果 V 不是由一有限集生成，则 V 称为**无穷维的**.

例 1 $\mathbb{R}^n$ 的标准基含有 n 个向量，所以 $\dim\mathbb{R}^n=n$，标准的多项式基 $\{1,t,t^2\}$ 表明 $\dim\mathbb{P}_2=3$，一般而言，$\dim\mathbb{P}_n=n+1$，所有多项式的空间 $\mathbb{P}$ 是无穷维的（见习题 27）. ■

例 2 令 $H=\mathrm{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$，这里 $\boldsymbol{v}_1=\begin{bmatrix}3\\6\\2\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}-1\\0\\1\end{bmatrix}$，则 H 是 4.4 节的例 7 中研究过的平面，H 的一个基为 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$，这是由于 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 不是倍数关系从而线性无关，于是 $\dim H=2$. ■

例 3 求下列子空间的维数：

$$H=\left\{\begin{bmatrix}a-3b+6c\\5a+4d\\b-2c-d\\5d\end{bmatrix}:a,b,c,d\in\mathbb{R}\right\}$$

解 易见 H 为下列向量的所有线性组合的集合：

$$\boldsymbol{v}_1=\begin{bmatrix}1\\5\\0\\0\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}-3\\0\\1\\0\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}6\\0\\-2\\0\end{bmatrix},\boldsymbol{v}_4=\begin{bmatrix}0\\4\\-1\\5\end{bmatrix}$$

显然 $\boldsymbol{v}_1\neq\boldsymbol{0}$. $\boldsymbol{v}_2$ 不是 $\boldsymbol{v}_1$ 的倍数，但 $\boldsymbol{v}_3$ 是 $\boldsymbol{v}_2$ 的倍数，由生成集定理，去掉 $\boldsymbol{v}_3$ 仍可生成 H. 最后由于 $\boldsymbol{v}_4$ 不是 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的线性组合，所以 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_4\}$ 是线性无关的（由 4.3 节中定理 4），进而是 H 的一个基，

㊀ 定理 9 也可应用到 V 的无限集中，一个无限集称为线性相关的，如果其中某一有限集是线性相关的，否则这个无限集是线性无关的. 若 S 是 V 中一个无限集，任取 S 的子集 $\{\boldsymbol{u}_1,\dots,\boldsymbol{u}_p\}$, $p>n$，上面的证明表明这个子集是线性相关的，从而 S 也是线性相关的.

于是 $\dim H = 3$. ■

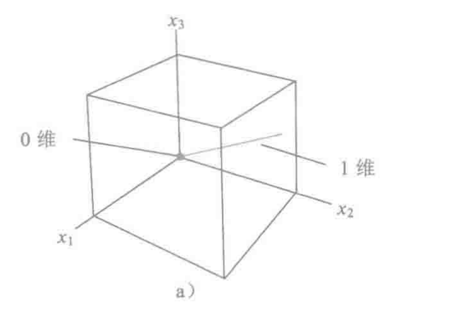
例 4 $\mathbb{R}^3$ 的子空间可用维数分类，见图 4-23.

0 维子空间. 只有零子空间是 0 维子空间.

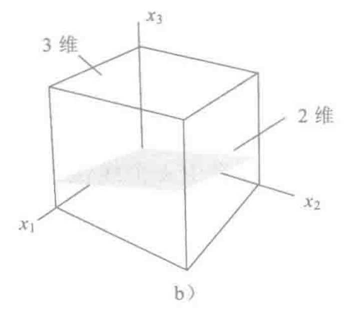
1 维子空间. 任一由单一非零向量生成的子空间，这样的子空间是经过原点的直线.

2 维子空间. 任一个由两个线性无关向量生成的子空间，这样的子空间是通过原点的平面.

3 维子空间. 只有 $\mathbb{R}^3$ 本身是 3 维子空间. 由可逆矩阵定理，$\mathbb{R}^3$ 中任意 3 个线性无关向量生成整个 $\mathbb{R}^3$.



a)



b)

图 4-23 $\mathbb{R}^3$ 的子空间样本 ■

有限维空间的子空间

下一个定理是生成集定理的一个自然配对.

定理 11 令 H 是有限维向量空间 V 的子空间，若有需要的话，H 中任一个线性无关集均可以扩充成为 H 的一个基，H 也是有限维的并且

$$\dim H \leqslant \dim V$$

证 若 $H=\{\mathbf{0}\}$，必然有 $\dim H = 0 \leqslant \dim V$. 否则，令 $S=\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_k\}$ 是 H 中任一线性无关集，若 S 生成 H，则 H 是 S 的一个基. 否则，存在 H 中某向量 $\boldsymbol{u}_{k+1}$ 不在 Span S 中，但 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_k,\boldsymbol{u}_{k+1}\}$ 将会是线性无关的，这是因为此集中没有一个向量可以表示为其前面向量的线性组合（由定理 4）.

只要这个新集合不能生成 H，我们就可以继续这个扩充 S 到 H 中一个更大的线性无关集的过程，但由定理 9，S 的线性无关扩充中向量的个数永远不能超过 V 的维数，所以 S 的扩充最终会生成 H 而且将成为 H 的一个基，同时 $\dim H \leqslant \dim V$. ■

当一个线性空间或子空间的维数知道后，通过下一个定理，求一个基就简单了. 即如果一个集合有适当个数的元素，则我们仅需要证明或者这个集合是线性无关的或者它生成这个空间. 这个定理在许多应用问题（例如包括微分方程和差分方程）均具有非常重要的意义，这里线性无关性比生成性更容易验证.

定理 12 （基定理）

令 V 是一个 p 维向量空间，$p\geqslant 1$，V 中任意含有 p 个元素的线性无关集必然是 V 的一个基.

任意含有 p 个元素且生成 V 的集合必然是 V 的一个基.

证 由定理 11，含 p 个元素的线性无关集 S 可以扩充为 V 的一个基. 但由于 $\dim V = p$，基必须恰好包含 p 个向量，所以 S 已经是 V 的一个基. 现假设 S 含有 p 个向量且生成 V，因为 V 是非零的，生成集定理蕴涵 S 的一个子集 S' 是 V 的一个基，因为 $\dim V = p, S'$ 一定包含 p 个向量，从而 $S = S'$. ■

Nul *A* 和 Col *A* 的维数

由于矩阵 A 的主元列构成 Col A 的一个基，我们一旦知道主元列，就知道了 Col A 的维数. 对 Nul A 的维数似乎需要做更多的工作，因为求 Nul A 的一个基通常比求 Col A 的一个基需要更多的时间，但有一条捷径可走！

令 A 为一个 $m \times n$ 矩阵，假设方程 $A\mathbf{x} = \mathbf{0}$ 有 k 个自由变量. 由 4.2 节，我们知道求 Nul A 的生成集的标准方法将恰好产生 k 个线性无关向量，比如说是 $\mathbf{u}_1, \cdots, \mathbf{u}_k$，每一个向量对应一个自由变量. 所以 $\{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$ 是 Nul A 的一个基，自由变量的个数决定基的大小. 为了后面的参考，让我们总结一下这些事实.

> Nul A 的维数是方程 $A\mathbf{x} = \mathbf{0}$ 中自由变量的个数，Col A 的维数是 A 中主元列的个数.

例 5 求 A 的零空间和列空间的维数.

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

解 将增广矩阵 $[A \quad \mathbf{0}]$ 行化简成阶梯形得

$$\begin{bmatrix} 1 & -2 & 2 & 3 & -1 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

有 3 个自由变量—— x_2, x_4 和 x_5，于是 Nul A 的维数是 3，由于 A 有两个主元列，所以 $\dim \text{Col}\, A = 2$. ■

练习题

判定下列每个命题的真假，对每个答案给出理由. 这里 V 是一非零有限维向量空间.

1. 若 $\dim V = p, S$ 是 V 的一个线性相关的子集，则 S 包含多于 p 个向量.
2. 若 S 生成 V，T 是 V 的一个子集且含有向量个数多于 S 中向量个数，则 T 是线性相关的.

习题 4.5

对习题 1~8 中的子空间，(a)求一个基，(b)说出维数.

1. $\left\{ \begin{bmatrix} s-2t \\ s+t \\ 3t \end{bmatrix} : s,t \in \mathbb{R} \right\}$

2. $\left\{ \begin{bmatrix} 2a \\ -4b \\ -2a \end{bmatrix} : a,b \in \mathbb{R} \right\}$

3. $\left\{ \begin{bmatrix} 2c \\ a-b \\ b-3c \\ a+2b \end{bmatrix} : a,b,c \in \mathbb{R} \right\}$

4. $\left\{ \begin{bmatrix} p+2q \\ -p \\ 3p-q \\ p+q \end{bmatrix} : p,q \in \mathbb{R} \right\}$

5. $\left\{\begin{bmatrix} p-2q \\ 2p+5r \\ -2q+2r \\ -3p+6r \end{bmatrix} : p,q,r \in \mathbb{R}\right\}$

6. $\left\{\begin{bmatrix} 3a-c \\ -b-3c \\ -7a+6b+5c \\ -3a+c \end{bmatrix} : a,b,c \in \mathbb{R}\right\}$

7. $\{(a,b,c): a-3b+c=0, b-2c=0, 2b-c=0\}$

8. $\{(a,b,c,d): a-3b+c=0\}$

9. 求 $\mathbb{R}^3$ 中所有第一个和第三个数字相等的全体向量生成的子空间的维数.

10. 求 $\mathbb{R}^2$ 中由 $\begin{bmatrix} 1 \\ -5 \end{bmatrix}$, $\begin{bmatrix} -2 \\ 10 \end{bmatrix}$, $\begin{bmatrix} -3 \\ 15 \end{bmatrix}$ 生成的子空间 H 的维数.

在习题 11 和习题 12 中，求已给向量生成的子空间的维数.

11. $\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, $\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 5 \\ 2 \\ 2 \end{bmatrix}$

12. $\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$, $\begin{bmatrix} -3 \\ -6 \\ 1 \end{bmatrix}$, $\begin{bmatrix} -2 \\ 3 \\ 5 \end{bmatrix}$, $\begin{bmatrix} -3 \\ 5 \\ 5 \end{bmatrix}$

确定习题 13~18 中给出矩阵的 Nul A 和 Col A 的维数.

13. $A=\begin{bmatrix} 1 & -6 & 9 & 0 & -2 \\ 0 & 1 & 2 & -4 & 5 \\ 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

14. $A=\begin{bmatrix} 1 & 2 & -4 & 3 & -2 & 6 & 0 \\ 0 & 0 & 0 & 1 & 0 & -3 & 7 \\ 0 & 0 & 0 & 0 & 1 & 4 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

15. $A=\begin{bmatrix} 1 & 2 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$

16. $A=\begin{bmatrix} 3 & 2 \\ -6 & 5 \end{bmatrix}$

17. $A=\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$

18. $A=\begin{bmatrix} 1 & 1 & -1 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

在习题 19 和习题 20 中，V 是一个向量空间，标出每个命题的真假，给出理由.

19. a. 矩阵的主元列的个数等于列空间的维数.
    b. $\mathbb{R}^3$ 中的平面是 $\mathbb{R}^3$ 的二维子空间.
    c. 向量空间 $\mathbb{P}_4$ 的维数为 4.
    d. 若 $\dim V = n$，S 是 V 中一个线性无关集，则 S 是 V 的一个基.
    e. 若集合 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 生成一个有限维向量空间 V，T 是 V 中多于 P 个向量的集合，则 T 是线性相关的.

20. a. $\mathbb{R}^2$ 是 $\mathbb{R}^3$ 的一个二维子空间.
    b. 方程 $A\boldsymbol{x}=\boldsymbol{0}$ 中变元个数等于 Nul A 的维数.
    c. 一个向量空间是无穷维的，如果它由一个无限集生成.
    d. 若 $\dim V = n$，S 生成 V，则 S 是 V 的一个基.
    e. $\mathbb{R}^3$ 只有一个三维子空间即 $\mathbb{R}^3$ 本身.

21. 前 4 个埃尔米特（Hermite）多项式为 $1, 2t, -2+4t^2$ 和 $-12t+8t^3$，这些多项式是在研究数学物理中的某种重要的微分方程时产生的.[⊖] 证明：这前 4 个埃尔米特多项式构成 $\mathbb{P}_3$ 的一个基.

22. 前 4 个拉盖尔（Laguerre）多项式为 $1, 1-t$, $2-4t+t^2$ 和 $6-18t+9t^2-t^3$. 证明：这些多项式构成 $\mathbb{P}_3$ 的一个基.

23. 令 $\mathcal{B}$ 是 $\mathbb{P}_3$ 的一个基，它由 21 题中埃尔米特多项式组成，令 $\boldsymbol{p}(t)=-1+8t^2+8t^3$，求 $\boldsymbol{p}$ 相对于 $\mathcal{B}$ 的坐标向量.

24. 令 $\mathcal{B}$ 是 $\mathbb{P}_2$ 的一个基，它由 22 题中前 3 个拉盖尔多项式组成，令 $\boldsymbol{p}(t)=5+5t-2t^2$，求 $\boldsymbol{p}$ 相对于 $\mathcal{B}$ 的坐标向量.

25. 令 S 是 n 维向量空间 V 的一个子集，设 S 包含

⊖ 参见 *Introduction to Functional Analysis*, 2d ed., A.E. Taylor and David C. Lay (New York: John Wiley & Sons, 1980), pp. 92-93. 此书还讨论了其他多项式.

少于 n 个向量，解释为什么 S 不能生成 V.

26. 令 H 是 n 维向量空间 V 的一个 n 维子空间，证明 $H=V$.

27. 解释为什么所有多项式的空间 $\mathbb{P}$ 是一个无穷维空间.

28. 证明：定义在实数直线上的全体连续函数的空间 $C(\mathbb{R})$ 是一个无穷维空间.

在习题 29 和习题 30 中，V 是一个非零有限维向量空间，列出的向量在 V 中，标出每个命题的真假，给出理由.（这些问题比习题 19 和习题 20 难些.）

29. a. 若存在集合 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 生成 V，则 $\dim V \leqslant p$.
    b. 若 V 中存在一个线性无关集 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$，则 $\dim V \geqslant p$.
    c. 若 $\dim V = p$，则 V 中存在一个 $p+1$ 个向量的生成集.

30. a. 若 V 中存在一线性相关集 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$，则 $\dim V \leqslant p$.
    b. 若 V 中任意由 p 个元素组成的集合均不能生成 V，则 $\dim V > p$.
    c. 若 $p \geqslant 2, \dim V = p$, 则每个由 $p-1$ 个向量构成的集合是线性无关的.

在习题 31 和习题 32 中，涉及有限维向量空间 V 和 W，还有线性变换 $T:V\to W$.

31. 令 H 是 V 的非零子空间，$T(H)$ 为 H 中向量的像集，则由 4.2 节中习题 35，$T(H)$ 是 W 的子空间. 证明 $\dim T(H) \leqslant \dim H$.

32. 令 H 是 V 的非零子空间，T 是一个由 V 到 W 内的一一（线性）映射，证明 $\dim T(H) = \dim H$. 若 T 是一个由 V 到 W 上的一一映射，则 $\dim V = \dim W$. 同构的有限维向量空间具有相同的维数.

33. [M] 按照定理 11，$\mathbb{R}^n$ 中一线性无关集 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k\}$ 可以扩充为 $\mathbb{R}^n$ 的一个基. 一个方法是构造矩阵 $A=[\boldsymbol{v}_1\cdots\boldsymbol{v}_k\ \ \boldsymbol{e}_1\cdots\boldsymbol{e}_n]$，其中 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n$ 是单位矩阵的列，那么 A 的主元列构成 $\mathbb{R}^n$ 的一个基.
    a. 利用这个方法将下列向量扩充为 $\mathbb{R}^5$ 的一个基.

$$\boldsymbol{v}_1=\begin{bmatrix}-9\\-7\\8\\-5\\7\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}9\\4\\1\\6\\-7\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}6\\7\\-8\\5\\-7\end{bmatrix}$$

    b. 解释下列问题：为什么原来的向量 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k$ 包含在 $\operatorname{Col} A$ 的基中？为什么 $\operatorname{Col} A=\mathbb{R}^n$？

34. [M] 令 $\mathcal{B}=\{1,\cos t,\cos^2 t,\cdots,\cos^6 t\}, \mathcal{C}=\{1,\cos t,\cos 2t,\cdots,\cos 6t\}$，假设有下列三角恒等式（见 4.1 节习题 37）：

$$\cos 2t=-1+2\cos^2 t$$
$$\cos 3t=-3\cos t+4\cos^3 t$$
$$\cos 4t=1-8\cos^2 t+8\cos^4 t$$
$$\cos 5t=5\cos t-20\cos^3 t+16\cos^5 t$$
$$\cos 6t=-1+18\cos^2 t-48\cos^4 t+32\cos^6 t$$

令 H 为 $\mathcal{B}$ 中函数生成的子空间，则由 4.3 节中习题 38，$\mathcal{B}$ 是 H 的一个基.
    a. 写出 $\mathcal{C}$ 中向量的 $\mathcal{B}$－坐标向量，利用它们证明 $\mathcal{C}$ 在 H 中是线性无关集.
    b. 解释为什么 $\mathcal{C}$ 是 H 的一个基.

练习题答案

1. 错，考虑集合 $\{\boldsymbol{0}\}$.
2. 对，由生成集定理，S 包含 V 的一个基，称这个基为 S'. 从而 T 包含比 S' 更多的向量，由定理 9，T 是线性相关的.

4.6 秩

由向量空间概念的帮助，本节我们观察矩阵的内部，揭示几个有趣且有用的隐藏在行和列之中的关系.

例如，设想在一个 40×50 矩阵 A 中放置 2 000 个随机数，然后确定 A 中线性无关列的最大个数和 A^{T} 中线性无关列的最大个数（即 A 中线性无关行的最大个数）. 值得注意的是，这两个数是相同的，我们将要看到，这个公共值是矩阵 A 的秩. 为了解释这个现象，需要检查由 A 的行生成的子空间.

行空间

若 A 是一个 $m\times n$ 矩阵，A 的每一行具有 n 个数字，即可以视为 $\mathbb{R}^n$ 中一个向量. 其行向量的所有线性组合的集合称为 A 的**行空间**，记为 Row A. 由于每一行具有 n 个数，所以 Row A 是 $\mathbb{R}^n$ 的一个子空间. 因为 A 的行与 A^{T} 的列相同，我们也可用 Col A^{T} 代替 Row A.

例 1 令

$$A=\begin{bmatrix}-2 & -5 & 8 & 0 & -17\\ 1 & 3 & -5 & 1 & 5\\ 3 & 11 & -19 & 7 & 1\\ 1 & 7 & -13 & 5 & -3\end{bmatrix}\qquad \begin{array}{l}\boldsymbol{r}_1=(-2,-5,8,0,-17)\\ \boldsymbol{r}_2=(1,3,-5,1,5)\\ \boldsymbol{r}_3=(3,11,-19,7,1)\\ \boldsymbol{r}_4=(1,7,-13,5,-3)\end{array}$$

A 的行空间是 $\mathbb{R}^5$ 的子空间，它由 $\{\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3,\boldsymbol{r}_4\}$ 生成. 即 Row $A=\mathrm{Span}(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3,\boldsymbol{r}_4)$，将行向量横着写是很自然的，然而，有时为了方便，它们也可以写成列向量. ■

在例 1 中，如果我们知道一些关于 A 的行之间的线性相关关系，那么我们可以利用生成集定理将这个生成集缩小到一个基. 不幸的是，在 A 上进行行变换对我们没有帮助，因为行变换改变了行的相关关系. 但行化简 A 当然是有价值的，正如下一个定理展示的那样！

定理 13 *若两个矩阵 A 和 B 行等价，则它们的行空间相同. 若 B 是阶梯形矩阵，则 B 的非零行构成 A 的行空间的一个基同时也是 B 的行空间的一个基.*

证 若 B 是由 A 进行行变换得到的，则 B 的行是 A 的行的线性组合，于是 B 的行的任意线性组合自然是 A 的行的线性组合. 从而 B 的行空间包含于 A 的行空间，因为行变换可逆，同理知 A 的行空间是 B 的行空间的子集，从而这两个空间相同. 若 B 是一个阶梯形，则其非零行是线性无关的，这是因为任何一个非零行均不为它下面的非零行的线性组合，于是 B 的非零行构成 B 的行空间的一个基，当然也是 A 的行空间的一个基. ■

本节的主要结论涉及三个空间：Row A, Col A 和 Nul A，下面的例子为这个结论做准备，同时展示如何通过对 A 进行的一系列行变换得到三个空间的基.

例 2 分别求矩阵 A 的行空间、列空间和零空间的基.

$$A=\begin{bmatrix}-2 & -5 & 8 & 0 & -17\\ 1 & 3 & -5 & 1 & 5\\ 3 & 11 & -19 & 7 & 1\\ 1 & 7 & -13 & 5 & -3\end{bmatrix}$$

解 为了求行空间和列空间的基，行化简 A 成阶梯形：

$$A\sim B=\begin{bmatrix}1 & 3 & -5 & 1 & 5\\ 0 & 1 & -2 & 2 & -7\\ 0 & 0 & 0 & -4 & 20\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}$$

由定理 13，B 的前 3 行构成 A 的行空间的一个基（也是 B 的行空间的基），从而

$$\text{Row } A\text{的基：}\ \{(1,3,-5,1,5),(0,1,-2,2,-7),(0,0,0,-4,20)\}$$

对列空间，观察 B，主元列在第 1，2 和 4 列，从而 A 的第 1，2 和 4 列（不是 B 的）构成 Col A 的一个基.

$$\text{Col } A\text{的基：}\ \left\{\begin{bmatrix}-2\\1\\3\\1\end{bmatrix},\begin{bmatrix}-5\\3\\11\\7\end{bmatrix},\begin{bmatrix}0\\1\\7\\5\end{bmatrix}\right\}$$

注意到 A 的任何阶梯形提供（在它的非零行）Row A 的一个基同时对 Col A 确定了 A 的主元列. 然而对 Nul A，则需要简化阶梯形，则 B 进一步行变换得：

$$A\sim B\sim C=\begin{bmatrix}1 & 0 & 1 & 0 & 1\\ 0 & 1 & -2 & 0 & 3\\ 0 & 0 & 0 & 1 & -5\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}$$

方程 $A\mathbf{x}=\mathbf{0}$ 等价于 $C\mathbf{x}=\mathbf{0}$，即

$$\begin{aligned}x_1+\ \ x_3+\ \ x_5&=0\\ x_2-2x_3+3x_5&=0\\ x_4-5x_5&=0\end{aligned}$$

所以 $x_1=-x_3-x_5, x_2=2x_3-3x_5, x_4=5x_5, x_3$ 和 x_5 为自由变量，通常的计算（在 4.2 节中讨论过）表明：

$$\text{Nul } A\text{的基：}\ \left\{\begin{bmatrix}-1\\2\\1\\0\\0\end{bmatrix},\begin{bmatrix}-1\\-3\\0\\5\\1\end{bmatrix}\right\}$$

通过观察可见，与 Col A 的基不同，Row A 和 Nul A 的基与 A 中的数没有简单的联系.[⊖] ■

⊖ 利用 A 的行求行空间 Row A 的基是可以办到的，对 A^T 进行行化简，直到 A^T 的主元列被求出，这些主元列是 A 的行，它们构成 A 的行空间的一个基.

警告 在例 2 中，尽管 B 的前 3 行是线性无关的，但由此推断说 A 的前 3 行是线性无关则是错误的（事实上 A 的第 3 行等于 2 倍的第 1 行加上 7 倍的第 2 行），行变换对矩阵的行不保持线性相关关系.

秩定理

下一定理描述 Col A，Row A 和 Nul A 的维数之间的基本关系.

定义 *A 的秩即 A 的列空间的维数.*

由于 Row A 与 Col A^{T} 相同，则 A 的行空间的维数等于 A^{T} 的秩. 零空间的维数有时被称为 A 的**零维**，尽管我们不使用这个术语.

做过 4.5 节的习题或读完上面的例 2 后，警觉的读者可能已经发现下一个定理的部分或全部结论.

定理 14 （秩定理）

$m\times n$ 矩阵 A 的列空间和行空间的维数相等，这个公共的维数（即 A 的秩）还等于 A 的主元位置的个数且满足方程

$$\operatorname{rank} A+\dim \operatorname{Nul} A=n$$

证 由 4.3 节定理 6，rank A 是 A 中主元列的个数，等价地，rank A 是 A 的阶梯形 B 中主元位置的个数. 进一步，因为 B 对每个主元有一个非零行，同时这些行对 A 的行空间而言构成一个基，所以 A 的秩也等于 A 的行空间的维数.

由 4.5 节，Nul A 的维数等于方程 $A\mathbf{x}=\mathbf{0}$ 中自由变量的个数，换句话说，Nul A 的维数是 A 中非主元列的个数，显然有

$$\{\text{主元列个数}\}+\{\text{非主元列个数}\}=\{\text{列的个数}\}$$

这就完成了证明. ■

定理 14 后面的思想在例 2 中的计算中可以看到. 阶梯形 B 的三个主元位置确定了基本变量同时确定了 Col A 和 Row A 的基向量.

例 3 a. 若 A 是一个 7×9 矩阵，具有二维零空间，A 的秩是多少？

b. 一个 6×9 矩阵能有二维零空间吗？

解 a. 由 A 有 9 列，$(\operatorname{rank} A)+2=9$，从而 $\operatorname{rank} A=7$.

b. 不能，若 B 为一个 6×9 矩阵，具有 2 维零空间，它的秩一定等于 7（由秩定理），但 B 的列是 $\mathbb{R}^6$ 中向量，从而 Col B 的维数不能超过 6，即 rank B 不能超过 6. ■

下一个例子为我们一直在研究的子空间直观化提供了一个好的方法，在第 6 章中，我们将学习到 Row A 和 Nul A 的公共向量只有零向量，并且事实上二者是相互“垂直”的，对 Row A^{T}(= Col A) 和 Nul A^{T} 有同样的结果. 所以例 4 中的图 4-24 对一般情形建立了一个良好的理解图像.（A^{T} 和 A 在一起研究的价值见习题 29 中.）

例 4 令 $A=\begin{bmatrix}3 & 0 & -1\\ 3 & 0 & -1\\ 4 & 0 & 5\end{bmatrix}$，容易检验 Nul A 是 x_2 轴，Row A 是 x_1x_3 平面，Col A 是方程为 $x_1-x_2=0$ 的平面. Nul A^{T} 是所有 $(1,-1,0)$ 的倍数构成的集合. 图 4-24 展示出在线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的定义域中的 Nul A 和 Row A；这个映射的值域 Col A 连同 Nul A^{T} 一起，被表现为 $\mathbb{R}^3$ 的一个分离的拷贝.

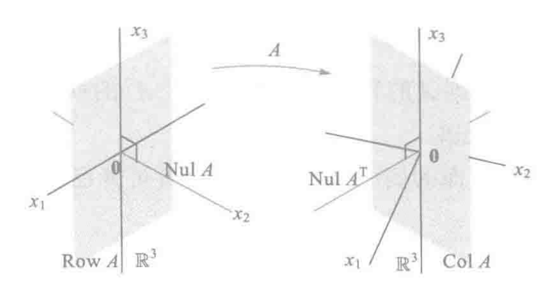


图 4-24 由矩阵 A 确定的子空间 ■

应用到方程组

秩定理对线性方程组的信息进行处理是一个有力的工具. 下一个例子模拟一个利用线性方程的现实问题，可能前面陈述过，例中没有直接用到线性代数的术语如矩阵、子空间、维数等.

例 5 一个科学家对一个 40 个方程 42 个变量的齐次方程组已经求出了两个解，这两个解不是倍数关系，而且其他所有解均能表示为这两个解的适当倍数之和. 这个科学家能确定一个相应的非齐次方程组（与此齐次方程组有相同的系数）有解吗？

解 有解，令 A 是这个方程组的 40×42 系数矩阵，已给的条件蕴涵这两个解是线性无关的且能生成 Nul A，所以 $\dim \mathrm{Nul}\, A=2$. 由秩定理，$\dim \mathrm{Col}\, A=42-2=40$. 由于 $\mathbb{R}^{40}$ 是 $\mathbb{R}^{40}$ 的唯一的维数是 40 的子空间，则 Col A 一定等于 $\mathbb{R}^{40}$，这表明每个非齐次方程 $A\boldsymbol{x}=\boldsymbol{b}$ 有一个解. ■

*秩和可逆矩阵定理

与矩阵相关的各种线性空间的概念对可逆矩阵定理提供了更多的一些命题，这里我们仅列出新的命题，但我们把它们接在 2.3 节中原来的可逆矩阵定理的后面.

定理 （可逆矩阵定理（续））

令 A 是一个 $n\times n$ 矩阵，则下列的命题中的每个均等价于 A 是可逆矩阵：

m. A 的列构成 $\mathbb{R}^n$ 的一个基.

n. Col $A=\mathbb{R}^n$.

o. $\dim \mathrm{Col}\, A=n$.

p. rank $A=n$.

q. Nul $A=\{\mathbf{0}\}$.

r. $\dim \mathrm{Nul}\, A=0$.

证 命题（m）从线性无关和生成的角度看，与命题（e）和（h）是逻辑上等价的. 至于上面其他五个命题，可由下列常见的蕴涵关系将它们与这个定理早期的一个命题链接起来：

$$(g)\Rightarrow(n)\Rightarrow(o)\Rightarrow(p)\Rightarrow(r)\Rightarrow(q)\Rightarrow(d)$$

命题（g）称对 $\mathbb{R}^n$ 中的每个 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 至少有一个解，由此可以推出（n），因为 Col A 实际上就是使方程 $A\boldsymbol{x}=\boldsymbol{b}$ 相容的所有 $\boldsymbol{b}$ 的集合. 蕴涵式(n)⇒(o)⇒(p)由维数和秩的定义可以推出. 如果 A 的秩等于 n，即 A 的列的个数，则由秩定理有 $\dim \operatorname{Nul} A=0$，也就是 $\operatorname{Nul} A=\{\boldsymbol{0}\}$. 于是(p)⇒(r)⇒(q). 而且由(q)可以推出方程 $A\boldsymbol{x}=\boldsymbol{0}$ 只有平凡解，即命题(d).因为已经知道命题(d)和(g)与 A 是可逆的命题是等价的，于是定理证毕. ■

我们没有将关于 A 的行空间的显然的命题添加到可逆矩阵定理中，这是因为行空间是 A^{T} 的列空间. 回顾可逆矩阵定理的（1）即 A 可逆当且仅当 A^{T} 可逆，从而可逆矩阵定理的每个命题也适合 A^{T}，这样做将定理的长度加倍，从而产生一串超过 30 个的命题！

数值计算的注解 本课程中讨论的许多算法对理解概念和进行简单的手工运算是有用的. 然而，这些算法通常不适合大规模的现实问题.

秩的确定是一个很好的例子. 将矩阵简化为阶梯形然后再数主元个数似乎很容易，但是除非在一个数字被明确指定的矩阵上执行精确的算法，否则行变换可能改变一个矩阵表面上的秩. 比如，若矩阵 $\begin{bmatrix}5 & 7\\ 5 & x\end{bmatrix}$ 中的 x 值没有被当作 7 精确地存储在计算机中，那么秩可能是 1 或 2，这依赖于计算机是否将 $x-7$ 当作零处理.

在实际应用中，矩阵 A 的有效秩常由 A 的奇异值分解来确定，这将在 7.4 节中讨论. 通过这种分解还可以可靠地求出 Col A、Row A、Nul A 和 Nul A^{T} 的基.

练习题

下列矩阵是行等价的.

$$A=\begin{bmatrix}2 & -1 & 1 & -6 & 8\\ 1 & -2 & -4 & 3 & -2\\ -7 & 8 & 10 & 3 & -10\\ 4 & -5 & -7 & 0 & 4\end{bmatrix}\quad B=\begin{bmatrix}1 & -2 & -4 & 3 & -2\\ 0 & 3 & 9 & -12 & 12\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}$$

1. 求 rank A 和 $\dim \operatorname{Nul} A$.
2. 求 Col A 和 Row A 的基.
3. 如果想要求 Nul A 的一个基，下一步要做什么？
4. A^{T} 的行阶梯形中有多少个主元列？

习题 4.6

在习题 1~4 中，假设 A 行等价于 B，不用计算，写出 rank A 和 $\dim \operatorname{Nul} A$，再求 Col A、Row A、Nul A 的基.

1. $A=\begin{bmatrix}1 & -4 & 9 & -7\\ -1 & 2 & -4 & 1\\ 5 & -6 & 10 & 7\end{bmatrix}, B=\begin{bmatrix}1 & 0 & -1 & 5\\ 0 & -2 & 5 & -6\\ 0 & 0 & 0 & 0\end{bmatrix}$

2. $A=\begin{bmatrix}1 & 3 & 4 & -1 & 2\\ 2 & 6 & 6 & 0 & -3\\ 3 & 9 & 3 & 6 & -3\\ 3 & 9 & 0 & 9 & 0\end{bmatrix}$,

$B=\begin{bmatrix}1 & 3 & 4 & -1 & 2\\ 0 & 0 & 1 & -1 & 1\\ 0 & 0 & 0 & 0 & -5\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}$

3. $A=\begin{bmatrix}2 & 6 & -6 & 6 & 3 & 6\\ -2 & -3 & 6 & -3 & 0 & -6\\ 4 & 9 & -12 & 9 & 3 & 12\\ -2 & 3 & 6 & 3 & 3 & -6\end{bmatrix}$,

$B=\begin{bmatrix}2 & 6 & -6 & 6 & 3 & 6\\ 0 & 3 & 0 & 3 & 3 & 0\\ 0 & 0 & 0 & 0 & 3 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\end{bmatrix}$

4. $A=\begin{bmatrix}1 & 1 & -2 & 0 & 1 & -2\\ 1 & 2 & -3 & 0 & -2 & -3\\ 1 & -1 & 0 & 0 & 1 & 6\\ 1 & -2 & 2 & 1 & -3 & 0\\ 1 & -2 & 1 & 0 & 2 & -1\end{bmatrix}$,

$B=\begin{bmatrix}1 & 1 & -2 & 0 & 1 & -2\\ 0 & 1 & -1 & 0 & -3 & -1\\ 0 & 0 & 1 & 1 & -13 & -1\\ 0 & 0 & 0 & 0 & 1 & -1\\ 0 & 0 & 0 & 0 & 0 & 1\end{bmatrix}$

5. 若一个 4×7 矩阵 A 秩为 3，求 $\dim \mathrm{Nul}\, A$，$\dim \mathrm{Row}\, A$ 和 $\mathrm{rank}\, A^{\mathrm{T}}$.
6. 若一个 7×5 矩阵 A 秩为 2，求 $\dim \mathrm{Nul}\, A$，$\dim \mathrm{Row}\, A$ 和 $\mathrm{rank}\, A^{\mathrm{T}}$.
7. 假设 4×7 矩阵 A 有 4 个主元列，$\mathrm{Col}\, A=\mathbb{R}^4$ 吗？$\mathrm{Nul}\, A=\mathbb{R}^3$ 吗？解释你的答案.
8. 假设 6×8 矩阵 A 有 4 个主元列，$\dim \mathrm{Nul}\, A$ 是多少？ $\mathrm{Col}\, A=\mathbb{R}^4$ 吗？解释你的答案.
9. 若 4×6 矩阵 A 的零空间是 3 维的，A 的列空间的维数是多少？ $\mathrm{Col}\, A=\mathbb{R}^3$ 吗？为什么？
10. 若 8×7 矩阵 A 的零空间是 5 维的，A 的列空间的维数是多少？
11. 若 8×5 矩阵 A 的零空间是 3 维的，A 的行空间的维数是多少？
12. 若 5×4 矩阵 A 的零空间是 2 维的，A 的行空间的维数是多少？
13. 若 A 是 7×5 矩阵，A 的秩最大可能为多少？若 A 是 5×7 矩阵，A 的秩最大可能为多少？解释你的回答.
14. 若 A 是 5×4 矩阵，A 的行空间的维数最大可能为多少？如果 A 是 4×5 矩阵，A 的行空间的维数最大可能为多少？解释你的回答.
15. 若 A 为一个 3×7 矩阵，$\mathrm{Nul}\, A$ 的最小可能的维数是多少？
16. 若 A 是一个 7×5 矩阵，$\mathrm{Nul}\, A$ 的最小可能的维数是多少？

在 17 和 18 题中，A 是一个 $m\times n$ 矩阵，标出每个命题的真假，检验每个答案.

17. a. A 的行空间与 A^{T} 的列空间相同.
    b. 若 B 是 A 的任一个阶梯形，B 有 3 个非零行，则 A 的前 3 行构成 $\mathrm{Row}\, A$ 的一个基.
    c. 矩阵 A 的行空间与列空间具有相同的维数，甚至 A 不是方阵，此结论也成立.
    d. A 的行空间的维数与零空间的维数之和等于 A 的行数.
    e. 在计算机上，行变换可能改变一个矩阵的表面上的秩.
18. a. 若 B 是 A 的任一阶梯形，则 B 的主元列构成 A 的列空间的一个基.
    b. 行变换保持 A 的行之间的线性相关关系.
    c. A 的零空间维数等于 A 的非主元列个数.
    d. A^{T} 的行空间与 A 的列空间相同.
    e. 若 A 与 B 行等价，则它们的行空间相同.
19. 假设一个含有 5 个线性方程 6 个未知数的齐次方程组的解都是一个非零解的倍数，对方程右边每个可能选取的常数，此方程组将必然会有一个解吗？解释之.
20. 假设一个含有 6 个线性方程 8 个未知数的非齐次方程组有一个解，具有两个自由变量，若改

变方程右边某常数，使新的方程组不相容可能吗？解释之.

21. 假设一个含有 9 个线性方程 10 个未知数的非齐次方程组对右边所有可能的常数均有解，相应的齐次方程组可以找到两个不成倍数的解吗？讨论之.
22. 一个含有 10 个方程 12 个未知数的齐次方程组的所有解为一个确定的非零解的倍数，这个结论可能吗？讨论之.
23. 一个含有 12 个线性方程 8 个未知数的齐次方程组有两个固定的没有倍数关系的解，同时所有其他解均是此二解的线性组合，所有解的集合可以由少于 12 个齐次线性方程的方程组刻画吗？若可以，有多少个方程？讨论之.
24. 一个含有 7 个方程 6 个未知数的非齐次方程组，对右边某些常数，该方程组有唯一解可能吗？对右边任一组常数，该方程组有唯一解可能吗？解释之.
25. 一个科学家解一个含有 10 个线性方程 12 个未知数的非齐次方程组，发现未知数中有 3 个是自由变量，该科学家是否可断定，若方程右边改变，新的非齐次方程组将有一个解？讨论之.
26. 在统计理论中，通常需要一个矩阵是满秩的，即矩阵的秩取到最大，对一个 $m\times n$ 矩阵，其中行数大于列数，解释为什么该矩阵满秩当且仅当它的列是线性无关的.

在习题 27~29 中，涉及一个 $m\times n$ 矩阵 A 和通常被称为由 A 确定的基本子空间.

27. $\text{Row } A, \text{Col } A, \text{Nul } A, \text{Row } A^{\mathrm{T}}, \text{Col } A^{\mathrm{T}}$ 和 $\text{Nul } A^{\mathrm{T}}$ 中哪一个子空间在 $\mathbb{R}^m$ 中，哪一个在 $\mathbb{R}^n$ 中？其中有多少不同的子空间？
28. 验证以下等式：
    a. $\dim \text{Row } A + \dim \text{Nul } A = n$　A 的列数
    b. $\dim \text{Col } A + \dim \text{Nul } A^{\mathrm{T}} = m$　A 的行数
29. 利用 28 题解释为什么方程 $A\boldsymbol{x}=\boldsymbol{b}$ 对任意 $\boldsymbol{b}\in\mathbb{R}^m$ 有解当且仅当方程 $A^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{0}$ 仅有平凡解.
30. 假设 A 是 $m\times n$ 矩阵，$\boldsymbol{b}$ 在 $\mathbb{R}^m$ 中，为使方程 $A\boldsymbol{x}=\boldsymbol{b}$ 相容，关于两个数 $\text{rank}\,[A\ \ \boldsymbol{b}]$ 和 $\text{rank}\,A$ 一定有什么关系？

秩为 1 的矩阵在某些计算以及本书的几处上下文的理论处理中都是重要的，其中包括第 7 章中的奇异值分解. 可以证明 $m\times n$ 矩阵 A 的秩为 1 当且仅当它是一个外积，即对某 $\boldsymbol{u}\in\mathbb{R}^m$，某 $\boldsymbol{v}\in\mathbb{R}^n$，有 $A=\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$.习题 31~33 暗示为什么这个性质是正确的.

31. 验证：如果 $\boldsymbol{u}=\begin{bmatrix}2\\-3\\5\end{bmatrix},\boldsymbol{v}=\begin{bmatrix}a\\b\\c\end{bmatrix}$，则 $\text{rank}\,\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}\leqslant 1$.
32. 令 $\boldsymbol{u}=\begin{bmatrix}1\\2\end{bmatrix}$，求 $\mathbb{R}^3$ 中向量 v，使得 $\begin{bmatrix}1&-3&4\\2&-6&8\end{bmatrix}=\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$.
33. 令 A 为任意一个 $\text{rank}\,A=1$ 的 2×3 矩阵，令 $\boldsymbol{u}$ 为 A 的第一列，且 $\boldsymbol{u}\neq\boldsymbol{0}$，解释为什么存在一个 $\mathbb{R}^3$ 中的向量 $\boldsymbol{v}$ 使得 $A=\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$. 若 A 的第一列为零，这个结论如何修改？
34. 令 A 为一个秩为 $r>0$ 的 $m\times n$ 矩阵，令 U 为 A 的一个阶梯形，解释为何存在一个可逆矩阵 E 使得 $A=EU$，利用这个因子分解将 A 写成 r 个秩为 1 的矩阵之和.（提示：参考 2.4 节定理 10.）
35. [M]令 $A=\begin{bmatrix}7&-9&-4&5&3&-3&-7\\-4&6&7&-2&-6&-5&5\\5&-7&-6&5&-6&2&8\\-3&5&8&-1&-7&-4&8\\6&-8&-5&4&4&9&3\end{bmatrix}$.
    a. 构造矩阵 C 和 N，它们的列分别为 $\text{Col } A$ 和 $\text{Nul } A$ 的基，再构造矩阵 R，它的行构成 $\text{Row } A$ 的一个基.
    b. 构造矩阵 M，它的列构成 $\text{Nul } A^{\mathrm{T}}$ 的一个基，构造矩阵 $S=[R^{\mathrm{T}}\ \ N]$，$T=[C\ \ M]$，解释为什么 S 和 T 应为方阵，验证 S 和 T 均为可逆的.
36. [M]随机取一个 6×7 整数值矩阵 A，并且它的

秩最大为 4，重复习题 35，生成 A 的一个方法是：随机建立一个 6×4 整数值矩阵 J 和一个 4×7 整数值矩阵 K，令 $A=JK$.（见本章后面的补充习题 12；阅读学习指导中关于矩阵生成程序的说明.）

37. [M]令 A 是 35 题中的矩阵，构造一个矩阵 C，使它的列是 A 的主元列，再构造一个矩阵 R，使它的行是 A 的简化阶梯形的非零行，计算 CR 并讨论你得到的结果.

38. [M]对三个随机取的 5×7 整数值矩阵 A，秩分别为 5，4 和 3，重复 37 题.对任意矩阵 A，作关于 CR 与 A 有何关系的一个猜想，证明你的猜想.

练习题答案

1. A 有两个主元列，所以 $\operatorname{rank} A=2$，又由于 A 共有 5 列，所以 $\dim \operatorname{Nul} A=5-2=3$.
2. A 的主元列是前两列，所以 $\operatorname{Col} A$ 的一个基为

$$\{\boldsymbol{a}_1,\boldsymbol{a}_2\}=\left\{\begin{bmatrix}2\\1\\-7\\4\end{bmatrix},\begin{bmatrix}-1\\-2\\8\\-5\end{bmatrix}\right\}$$

B 的非零行构成 $\operatorname{Row} A$ 的一个基，即 $\{(1,-2,-4,3,-2),(0,3,9,-12,12)\}$，在此特殊例子中，$A$ 的任意两行构成行空间的一个基，这是因为行空间是 2 维的，同时 A 中任一行均不是另一行的倍数. 一般而言，A 的阶梯形的非零行通常可作为 $\operatorname{Row} A$ 的一个基，A 本身的行则不行.

3. 对 $\operatorname{Nul} A$，下一步是对 B 进行行变换得到 A 的简化阶梯形.
4. 由秩定理，因 $\operatorname{Col} A^{\mathrm{T}}=\operatorname{Row} A$,所以 $\operatorname{rank} A^{\mathrm{T}}=\operatorname{rank} A$，从而 A^{T} 有两个主元位置.

4.7 基的变换

对一个 n 维向量空间 V，当一个基 $\mathcal{B}$ 取定后，与之相关的映射到 $\mathbb{R}^n$ 上的坐标映射对 V 提供了一个坐标系. V 中每个向量 $\boldsymbol{x}$ 由它的 $\mathcal{B}-$坐标向量 $[\boldsymbol{x}]_{\mathcal{B}}$ 唯一确定.[㊀]

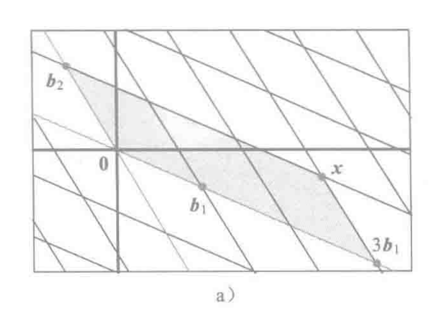
在某些应用中，一个问题开始是用一个基 $\mathcal{B}$ 描述，但问题的解可通过将 $\mathcal{B}$ 变为一个新的基 $\mathcal{C}$ 得到帮助（例子将在第 5 章和第 7 章中给出）. 每个向量被确定为一个新的 $\mathcal{C}-$坐标向量. 在本节中，我们研究对每个 $\boldsymbol{x}\in V$，$[\boldsymbol{x}]_{\mathcal{C}}$ 与 $[\boldsymbol{x}]_{\mathcal{B}}$ 如何联系起来.

为使问题直观化，考虑图 4-25 中的两个坐标系，在图 4-25a 中，$\boldsymbol{x}=3\boldsymbol{b}_1+\boldsymbol{b}_2$，同时在图 4-25b 中，同样的 $\boldsymbol{x}$ 表示为 $\boldsymbol{x}=6\boldsymbol{c}_1+4\boldsymbol{c}_2$，即

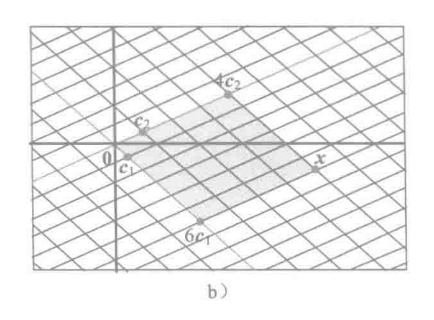
$$[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}3\\1\end{bmatrix},\ [\boldsymbol{x}]_{\mathcal{C}}=\begin{bmatrix}6\\4\end{bmatrix}$$

我们的问题是找到这两个坐标向量之间的联系，例 1 表明如何作这件事，使我们知道如何由 $\boldsymbol{c}_1$ 和 $\boldsymbol{c}_2$ 得到 $\boldsymbol{b}_1$ 和 $\boldsymbol{b}_2$.

㊀ 将 $[\boldsymbol{x}]_{\mathcal{B}}$ 看作 $\boldsymbol{x}$ 的一个“名字”，它给出权的列表将 $\boldsymbol{x}$ 表示为 $\mathcal{B}$ 中基向量的线性组合.



a)



b)

图 4-25　同样向量空间的两个坐标系

例 1　对一个向量空间 V，考虑两个基 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$ 和 $\mathcal{C}=\{\boldsymbol{c}_1,\boldsymbol{c}_2\}$，满足

$$\boldsymbol{b}_1=4\boldsymbol{c}_1+\boldsymbol{c}_2,\ \boldsymbol{b}_2=-6\boldsymbol{c}_1+\boldsymbol{c}_2 \tag{1}$$

假设

$$\boldsymbol{x}=3\boldsymbol{b}_1+\boldsymbol{b}_2 \tag{2}$$

即假设 $[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}3\\1\end{bmatrix}$，求 $[\boldsymbol{x}]_{\mathcal{C}}$.

解　对（2）中 $\boldsymbol{x}$ 应用由 $\mathcal{C}$ 确定的坐标映射，因为坐标映射是一个线性变换.

$$\begin{aligned}[\boldsymbol{x}]_{\mathcal{C}}&=[3\boldsymbol{b}_1+\boldsymbol{b}_2]_{\mathcal{C}}\\&=3[\boldsymbol{b}_1]_{\mathcal{C}}+[\boldsymbol{b}_2]_{\mathcal{C}}\end{aligned}$$

利用将线性组合中的向量看做矩阵的列，我们可以将这个向量方程写成一个矩阵方程：

$$[\boldsymbol{x}]_{\mathcal{C}}=[\,[\boldsymbol{b}_1]_{\mathcal{C}}\ \ [\boldsymbol{b}_2]_{\mathcal{C}}\,]\begin{bmatrix}3\\1\end{bmatrix} \tag{3}$$

只要我们知道该矩阵的列，这个公式就给出了 $[\boldsymbol{x}]_{\mathcal{C}}$，由（1），

$$[\boldsymbol{b}_1]_{\mathcal{C}}=\begin{bmatrix}4\\1\end{bmatrix},\ [\boldsymbol{b}_2]_{\mathcal{C}}=\begin{bmatrix}-6\\1\end{bmatrix}$$

于是（3）给出了解：

$$[\boldsymbol{x}]_{\mathcal{C}}=\begin{bmatrix}4&-6\\1&1\end{bmatrix}\begin{bmatrix}3\\1\end{bmatrix}=\begin{bmatrix}6\\4\end{bmatrix}$$

即与图 4-25 中 $\boldsymbol{x}$ 相匹配的 $\boldsymbol{x}$ 的 $\mathcal{C}-$坐标. ■

可以将推导出公式（3）的论证推广从而产生下列结果.（见习题 15 和习题 16.）

定理 15　设 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ 和 $\mathcal{C}=\{\boldsymbol{c}_1,\cdots,\boldsymbol{c}_n\}$ 是向量空间 V 的基，则存在一个 $n\times n$ 矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 使得

$$[\boldsymbol{x}]_{\mathcal{C}}=\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\boldsymbol{x}]_{\mathcal{B}} \tag{4}$$

$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的列是基 $\mathcal{B}$ 中向量的 $\mathcal{C}-$坐标向量，即

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}=[\,[\boldsymbol{b}_1]_{\mathcal{C}}\ \ [\boldsymbol{b}_2]_{\mathcal{C}}\ \ \cdots\ \ [\boldsymbol{b}_n]_{\mathcal{C}}\,] \tag{5}$$

定理 15 中矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 称为**由 $\mathcal{B}$ 到 $\mathcal{C}$ 的坐标变换矩阵**. 乘以 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的运算将 $\mathcal{B}$ – 坐标变为 $\mathcal{C}$ – 坐标，[㊀] 图 4-26 给出坐标变换方程（4）的说明.

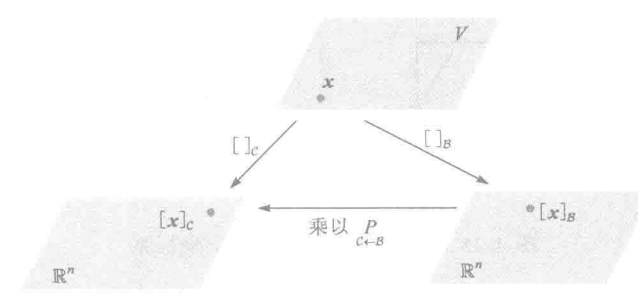


图 4-26　V 的两个坐标系

$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的列是线性无关的，这是因为它们是线性无关集 $\mathcal{B}$ 的坐标向量（见 4.4 节习题 25），于是得到 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 是可逆的. 将（4）两边左乘以 $\left(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}\right)^{-1}$，得

$$\left(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}\right)^{-1}[\boldsymbol{x}]_{\mathcal{C}}=[\boldsymbol{x}]_{\mathcal{B}}$$

于是 $\left(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}\right)^{-1}$ 是将 $\mathcal{C}$ – 坐标变为 $\mathcal{B}$ – 坐标的矩阵，即

$$\left(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}\right)^{-1}=\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} \tag{6}$$

$\mathbb{R}^n$ 中基的变换

若 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$，$\mathcal{E}$ 是 $\mathbb{R}^n$ 中的标准基 $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$，则 $[\boldsymbol{b}_1]_{\mathcal{E}}=\boldsymbol{b}_1$，$\mathcal{B}$ 中其他向量也类似. 在此情形下，$\underset{\mathcal{E}\leftarrow\mathcal{B}}{P}$ 与 4.4 节中引入的坐标变换矩阵 $P_{\mathcal{B}}$ 相同，即

$$P_{\mathcal{B}}=[\boldsymbol{b}_1\quad \boldsymbol{b}_2\quad \cdots\quad \boldsymbol{b}_n]$$

为了在 $\mathbb{R}^n$ 中两个非标准基之间变换坐标，我们需要定理 15，定理 15 表明为解决基变换问题，我们需要原来的基关于新的基的坐标向量.

例 2　设 $\boldsymbol{b}_1=\begin{bmatrix}-9\\1\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}-5\\-1\end{bmatrix},\boldsymbol{c}_1=\begin{bmatrix}1\\-4\end{bmatrix},\boldsymbol{c}_2=\begin{bmatrix}3\\-5\end{bmatrix}$，考虑 $\mathbb{R}^2$ 中基 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$，$\mathcal{C}=\{\boldsymbol{c}_1,\boldsymbol{c}_2\}$，求由 $\mathcal{B}$ 到 $\mathcal{C}$ 的坐标变换矩阵.

解　矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 涉及 $\boldsymbol{b}_1$ 和 $\boldsymbol{b}_2$ 的 $\mathcal{C}$ – 坐标向量，设 $[\boldsymbol{b}_1]_{\mathcal{C}}=\begin{bmatrix}x_1\\x_2\end{bmatrix},[\boldsymbol{b}_2]_{\mathcal{C}}=\begin{bmatrix}y_1\\y_2\end{bmatrix}$，于是由定义

$$[\boldsymbol{c}_1\quad \boldsymbol{c}_2]\begin{bmatrix}x_1\\x_2\end{bmatrix}=\boldsymbol{b}_1,\ [\boldsymbol{c}_1\quad \boldsymbol{c}_2]\begin{bmatrix}y_1\\y_2\end{bmatrix}=\boldsymbol{b}_2$$

㊀ 为了记住如何构造这个矩阵，将 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\boldsymbol{x}]_{\mathcal{B}}$ 看作 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的列的线性组合，这个矩阵－向量积是一个 $\mathcal{C}$ – 坐标向量，所以 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的列也应该是 $\mathcal{C}$ – 坐标向量.

为了同步解出这两个方程组，将 $\boldsymbol{b}_1$ 和 $\boldsymbol{b}_2$ 扩大到系数矩阵中并行化简之：

$$[\boldsymbol{c}_1\ \ \boldsymbol{c}_2\ \vdots\ \boldsymbol{b}_1\ \ \boldsymbol{b}_2]=\begin{bmatrix}1 & 3 & \vdots & -9 & -5\\ -4 & -5 & \vdots & 1 & -1\end{bmatrix}\sim\begin{bmatrix}1 & 0 & \vdots & 6 & 4\\ 0 & 1 & \vdots & -5 & -3\end{bmatrix} \tag{7}$$

于是 $[\boldsymbol{b}_1]_{\mathcal{C}}=\begin{bmatrix}6\\-5\end{bmatrix}$, $[\boldsymbol{b}_2]_{\mathcal{C}}=\begin{bmatrix}4\\-3\end{bmatrix}$. 因此所要求的坐标变换矩阵是

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}=[[\boldsymbol{b}_1]_{\mathcal{C}}\ \ [\boldsymbol{b}_2]_{\mathcal{C}}]=\begin{bmatrix}6 & 4\\ -5 & -3\end{bmatrix}$$ ■

观察例 2 中的矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$，它已经出现在（7）中，这并不会令人感到意外，因为 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的第 1 列是行化简 $[\boldsymbol{c}_1\ \ \boldsymbol{c}_2\ \vdots\ \boldsymbol{b}_1]$ 到 $[I\ \vdots\ [\boldsymbol{b}_1]_{\mathcal{C}}]$ 的结果，对 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的第 2 列也是类似的，于是

$$[\boldsymbol{c}_1\ \ \boldsymbol{c}_2\ \vdots\ \boldsymbol{b}_1\ \ \boldsymbol{b}_2]\sim\left[I\ \vdots\ \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}\right]$$

求 $\mathbb{R}^n$ 中的任意两个基之间的坐标变换矩阵具有类似的步骤.

例 3 设 $\boldsymbol{b}_1=\begin{bmatrix}1\\-3\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}-2\\4\end{bmatrix},\boldsymbol{c}_1=\begin{bmatrix}-7\\9\end{bmatrix},\boldsymbol{c}_2=\begin{bmatrix}-5\\7\end{bmatrix}$，考虑 $\mathbb{R}^2$ 中的基 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\},\mathcal{C}=\{\boldsymbol{c}_1,\boldsymbol{c}_2\}$.

a. 求由 $\mathcal{C}$ 到 $\mathcal{B}$ 的坐标变换矩阵.

b. 求由 $\mathcal{B}$ 到 $\mathcal{C}$ 的坐标变换矩阵.

解 a. 注意到求 $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$ 比求 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 更方便，计算

$$[\boldsymbol{b}_1\ \ \boldsymbol{b}_2\ \vdots\ \boldsymbol{c}_1\ \ \boldsymbol{c}_2]=\begin{bmatrix}1 & -2 & \vdots & -7 & -5\\ -3 & 4 & \vdots & 9 & 7\end{bmatrix}\sim\begin{bmatrix}1 & 0 & \vdots & 5 & 3\\ 0 & 1 & \vdots & 6 & 4\end{bmatrix}$$

所以

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}=\begin{bmatrix}5 & 3\\ 6 & 4\end{bmatrix}$$

b. 由（a）和上面的（5）式（将 $\mathcal{B}$ 和 $\mathcal{C}$ 互换）

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}=\left(\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}\right)^{-1}=\frac{1}{2}\begin{bmatrix}4 & -3\\ -6 & 5\end{bmatrix}=\begin{bmatrix}2 & -3/2\\ -3 & 5/2\end{bmatrix}$$ ■

另一个关于坐标变换矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的描述，是使用坐标变换矩阵 $P_{\mathcal{B}}$ 和 $P_{\mathcal{C}}$ 分别将 $\mathcal{B}$－坐标和 $\mathcal{C}$－坐标转换成为标准坐标. 回忆对 $\mathbb{R}^n$ 中的每个 $\boldsymbol{x}$，有

$$P_{\mathcal{B}}[\boldsymbol{x}]_{\mathcal{B}}=\boldsymbol{x},\ \ P_{\mathcal{C}}[\boldsymbol{x}]_{\mathcal{C}}=\boldsymbol{x},\ \ [\boldsymbol{x}]_{\mathcal{C}}=P_{\mathcal{C}}^{-1}\boldsymbol{x}$$

于是

$$[\boldsymbol{x}]_{\mathcal{C}}=P_{\mathcal{C}}^{-1}\boldsymbol{x}=P_{\mathcal{C}}^{-1}P_{\mathcal{B}}[\boldsymbol{x}]_{\mathcal{B}}$$

在 $\mathbb{R}^n$ 中，坐标变换矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 可以用 $P_{\mathcal{C}}^{-1}P_{\mathcal{B}}$ 来计算. 事实上，对于比 2×2 更大的矩阵，一个类似于例 3 的算法比先计算 $P_{\mathcal{C}}^{-1}$ 再计算 $P_{\mathcal{C}}^{-1}P_{\mathcal{B}}$ 的方法更快，见 2.2 节习题 12.

练习题

1. 设 $\mathcal{F}=\{\boldsymbol{f}_1,\boldsymbol{f}_2\},\mathcal{G}=\{\boldsymbol{g}_1,\boldsymbol{g}_2\}$ 为向量空间 V 的两个基，P 为一个矩阵，它的列是 $[\boldsymbol{f}_1]_{\mathcal{G}}$ 和 $[\boldsymbol{f}_2]_{\mathcal{G}}$，对所有 $\boldsymbol{v}\in V,P$ 满足下列哪一个方程？

(i) $[\boldsymbol{v}]_{\mathcal{F}}=P[\boldsymbol{v}]_{\mathcal{G}}$　　　(ii) $[\boldsymbol{v}]_{\mathcal{G}}=P[\boldsymbol{v}]_{\mathcal{F}}$

2. 设 $\mathcal{B}$ 和 $\mathcal{C}$ 如例 1 所示，利用例 1 的结果求由 $\mathcal{C}$ 到 $\mathcal{B}$ 的坐标变换矩阵.

习题 4.7

1. 设 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$ 和 $\mathcal{C}=\{\boldsymbol{c}_1,\boldsymbol{c}_2\}$ 是向量空间 V 的两个基，设

$$\boldsymbol{b}_1=6\boldsymbol{c}_1-2\boldsymbol{c}_2, \boldsymbol{b}_2=9\boldsymbol{c}_1-4\boldsymbol{c}_2$$

   a. 求由 $\mathcal{B}$ 到 $\mathcal{C}$ 的坐标变换矩阵.

   b. 利用（a），对 $\boldsymbol{x}=-3\boldsymbol{b}_1+2\boldsymbol{b}_2$，求 $[\boldsymbol{x}]_{\mathcal{C}}$.

2. 设 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$ 和 $\mathcal{C}=\{\boldsymbol{c}_1,\boldsymbol{c}_2\}$ 是向量空间 V 的两个基，设

$$\boldsymbol{b}_1=-2\boldsymbol{c}_1+4\boldsymbol{c}_2, \boldsymbol{b}_2=3\boldsymbol{c}_1-6\boldsymbol{c}_2$$

   a. 求由 $\mathcal{B}$ 到 $\mathcal{C}$ 的坐标变换矩阵.

   b. 对 $\boldsymbol{x}=2\boldsymbol{b}_1+3\boldsymbol{b}_2$，求 $[\boldsymbol{x}]_{\mathcal{C}}$.

3. 设 $\mathcal{U}=\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$ 和 $\mathcal{W}=\{\boldsymbol{w}_1,\boldsymbol{w}_2\}$ 是 V 的两个基，P 为一个矩阵，它的列为 $[\boldsymbol{u}_1]_{\mathcal{W}}$ 和 $[\boldsymbol{u}_2]_{\mathcal{W}}$. 对所有 $\boldsymbol{x}\in V$, P 满足以下哪一个方程？

   (i) $[\boldsymbol{x}]_{\mathcal{U}}=P[\boldsymbol{x}]_{\mathcal{W}}$　(ii) $[\boldsymbol{x}]_{\mathcal{W}}=P[\boldsymbol{x}]_{\mathcal{U}}$

4. 设 $\mathcal{A}=\{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3\}$ 和 $\mathcal{D}=\{\boldsymbol{d}_1,\boldsymbol{d}_2,\boldsymbol{d}_3\}$ 是 V 的两个基，$P=[[\boldsymbol{d}_1]_{\mathcal{A}}\ [\boldsymbol{d}_2]_{\mathcal{A}}\ [\boldsymbol{d}_3]_{\mathcal{A}}]$，对所有 $\boldsymbol{x}\in V, P$ 满足以下哪一个方程？

   (i) $[\boldsymbol{x}]_{\mathcal{A}}=P[\boldsymbol{x}]_{\mathcal{D}}$　(ii) $[\boldsymbol{x}]_{\mathcal{D}}=P[\boldsymbol{x}]_{\mathcal{A}}$

5. 设 $\mathcal{A}=\{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3\}$ 和 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3\}$ 是向量空间 V 的两个基，设 $\boldsymbol{a}_1=4\boldsymbol{b}_1-\boldsymbol{b}_2$，$\boldsymbol{a}_2=-\boldsymbol{b}_1+\boldsymbol{b}_2+\boldsymbol{b}_3$，$\boldsymbol{a}_3=\boldsymbol{b}_2-2\boldsymbol{b}_3$.

   a. 求由 $\mathcal{A}$ 到 $\mathcal{B}$ 的坐标变换矩阵.

   b. 对 $\boldsymbol{x}=3\boldsymbol{a}_1+4\boldsymbol{a}_2+\boldsymbol{a}_3$，求 $[\boldsymbol{x}]_{\mathcal{B}}$.

6. 设 $\mathcal{D}=\{\boldsymbol{d}_1,\boldsymbol{d}_2,\boldsymbol{d}_3\}$ 和 $\mathcal{F}=\{\boldsymbol{f}_1,\boldsymbol{f}_2,\boldsymbol{f}_3\}$ 是向量空间 V 的两个基，设 $\boldsymbol{f}_1=2\boldsymbol{d}_1-\boldsymbol{d}_2+\boldsymbol{d}_3$，$\boldsymbol{f}_2=3\boldsymbol{d}_2+\boldsymbol{d}_3$，$\boldsymbol{f}_3=-3\boldsymbol{d}_1+2\boldsymbol{d}_3$.

   a. 求由 $\mathcal{F}$ 到 $\mathcal{D}$ 的坐标变换矩阵.

   b. 对 $\boldsymbol{x}=\boldsymbol{f}_1-2\boldsymbol{f}_2+2\boldsymbol{f}_3$，求 $[\boldsymbol{x}]_{\mathcal{D}}$.

在习题 7~10 中，设 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$ 和 $\mathcal{C}=\{\boldsymbol{c}_1,\boldsymbol{c}_2\}$ 是 $\mathbb{R}^2$ 的两个基，求由 $\mathcal{B}$ 到 $\mathcal{C}$ 的坐标变换矩阵和由 $\mathcal{C}$ 到 $\mathcal{B}$ 的坐标变换矩阵.

7. $\boldsymbol{b}_1=\begin{bmatrix}7\\5\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}-3\\-1\end{bmatrix},\boldsymbol{c}_1=\begin{bmatrix}1\\-5\end{bmatrix},\boldsymbol{c}_2=\begin{bmatrix}-2\\2\end{bmatrix}$

8. $\boldsymbol{b}_1=\begin{bmatrix}-1\\8\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}1\\-7\end{bmatrix},\boldsymbol{c}_1=\begin{bmatrix}1\\2\end{bmatrix},\boldsymbol{c}_2=\begin{bmatrix}1\\1\end{bmatrix}$

9. $\boldsymbol{b}_1=\begin{bmatrix}4\\4\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}8\\4\end{bmatrix},\boldsymbol{c}_1=\begin{bmatrix}2\\2\end{bmatrix},\boldsymbol{c}_2=\begin{bmatrix}-2\\2\end{bmatrix}$

10. $\boldsymbol{b}_1=\begin{bmatrix}6\\-12\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}4\\2\end{bmatrix},\boldsymbol{c}_1=\begin{bmatrix}4\\2\end{bmatrix},\boldsymbol{c}_2=\begin{bmatrix}3\\9\end{bmatrix}$

在习题 11 和习题 12 中，$\mathcal{B}$ 和 $\mathcal{C}$ 是向量空间 V 的两个基，标出每个命题的真假，给出理由.

11. a. 坐标变换矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的列是 $\mathcal{C}$ 中向量的 $\mathcal{B}-$坐标向量.

    b. 若 $V=\mathbb{R}^n$，$\mathcal{C}$ 为 V 的标准基，则 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 与 4.4 节中引入的坐标变换矩阵 $P_{\mathcal{B}}$ 相同.

12. a. $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的列是线性无关的.

    b. 若 $V=\mathbb{R}^2,\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$ 和 $\mathcal{C}=\{\boldsymbol{c}_1,\boldsymbol{c}_2\}$，则将 $[\boldsymbol{c}_1\ \ \boldsymbol{c}_2\ \ \boldsymbol{b}_1\ \ \boldsymbol{b}_2]$ 行化简为 $[I\ \ P]$ 时产生一矩阵 P，满足对任意 $\boldsymbol{x}\in V$, 均有 $[\boldsymbol{x}]_{\mathcal{B}}=P[\boldsymbol{x}]_{\mathcal{C}}$.

13. 在 $\mathbb{P}_2$ 中，求由基 $\mathcal{B}=\{1-2t+t^2,3-5t+4t^2,2t+3t^2\}$ 到标准基 $\mathcal{C}=\{1,t,t^2\}$ 的坐标变换矩阵，再求 $-1+2t$ 的 $\mathcal{B}-$坐标向量.

14. 在 $\mathbb{P}_2$ 中，求由基 $\mathcal{B}=\{1-3t^2,2+t-5t^2,1+2t\}$ 到标准基的坐标变换矩阵，再将 t^2 写成 $\mathcal{B}$ 中多项式的线性组合.

在习题 15 和习题 16 中，通过填写正确的理由完成定理 15 的证明.

15. 给定 $\boldsymbol{v}\in V$，则对某些数 $x_1,\cdots,x_n$ 有

$$\boldsymbol{v}=x_1\boldsymbol{b}_1+x_2\boldsymbol{b}_2+\cdots+x_n\boldsymbol{b}_n$$

这是因为___(a)___，应用由基 $\mathcal{C}$ 确定的坐标映射，有

$$[\boldsymbol{v}]_{\mathcal{C}}=x_1[\boldsymbol{b}_1]_{\mathcal{C}}+x_2[\boldsymbol{b}_2]_{\mathcal{C}}+\cdots+x_n[\boldsymbol{b}_n]_{\mathcal{C}}$$

这是因为___(b)___，由定义___(c)___，我们可以将这个方程写成以下形式：

$$[\boldsymbol{v}]_{\mathcal{C}}=[\,[\boldsymbol{b}_1]_{\mathcal{C}}\ \ [\boldsymbol{b}_2]_{\mathcal{C}}\ \cdots\ [\boldsymbol{b}_n]_{\mathcal{C}}\,]\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} \quad (8)$$

因为（8）式右边的向量是___(d)___，这就证明了对于 $\boldsymbol{v}\in V$，（5）中的矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 满足 $[\boldsymbol{v}]_{\mathcal{C}}=\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\boldsymbol{v}]_{\mathcal{B}}$.

16. 设 Q 是任意矩阵，使得

$$[\boldsymbol{v}]_{\mathcal{C}}=Q[\boldsymbol{v}]_{\mathcal{B}} \quad 对\ \boldsymbol{v}\in V \quad (9)$$

在（9）中，设 $\boldsymbol{v}=\boldsymbol{b}_1$，则（9）表明 $[\boldsymbol{b}_1]_{\mathcal{C}}$ 是 Q 的第一列，因为___(a)___. 类似地，对 k=2，…，n，Q 的第 k 列是___(b)___，因为___(c)___. 这表明由定理 15 中的（5）定义的矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 是满足条件（4）的唯一矩阵.

17. [M] 设 $\mathcal{B}=\{\boldsymbol{x}_0,\cdots,\boldsymbol{x}_6\},\mathcal{C}=\{\boldsymbol{y}_0,\cdots,\boldsymbol{y}_6\}$，这里 $\boldsymbol{x}_k$ 是函数 $\cos^k t$，$\boldsymbol{y}_k$ 是函数 $\cos kt$，4.5 节习题 34 证明 $\mathcal{B}$ 和 $\mathcal{C}$ 都是向量空间 $H=\text{Span}\{\boldsymbol{x}_0,\cdots,\boldsymbol{x}_6\}$ 的基.
   a. 设 $P=[\,[\boldsymbol{y}_0]_{\mathcal{B}}\ \cdots\ [\boldsymbol{y}_6]_{\mathcal{B}}\,]$，计算 P^{-1}.
   b. 解释为什么 P^{-1} 的列是 $\boldsymbol{x}_0,\cdots,\boldsymbol{x}_6$ 的 $\mathcal{C}$－坐标向量，然后利用这些坐标向量写出表示 $\mathcal{C}$ 中函数 $\cos t$ 的幂的三角恒等式.

18. [M]（需要微积分的知识）[㊀]回顾微积分中如下积分：

$$\int(5\cos^3 t-6\cos^4 t+5\cos^5 t-12\cos^6 t)\mathrm{d}t \quad (10)$$

其计算过程是繁琐的.（通常的方法是重复利用分部积分法和半角公式.）使用 17 题的矩阵 P 或 P^{-1} 变换（10）式，再计算这个积分.

19. [M] 设 $P=\begin{bmatrix}1&2&-1\\-3&-5&0\\4&6&1\end{bmatrix}$，$\boldsymbol{v}_1=\begin{bmatrix}-2\\2\\3\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}-8\\5\\2\end{bmatrix}$，$\boldsymbol{v}_3=\begin{bmatrix}-7\\2\\6\end{bmatrix}$.
   a. 求 $\mathbb{R}^3$ 中的一个基 $\{\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3\}$，使得 P 是由基 $\{\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3\}$ 到 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 的坐标变换矩阵.（提示：$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的列代表什么？）
   b. 求 $\mathbb{R}^3$ 的一个基 $\{\boldsymbol{w}_1,\boldsymbol{w}_2,\boldsymbol{w}_3\}$，使得 P 是由基 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 到基 $\{\boldsymbol{w}_1,\boldsymbol{w}_2,\boldsymbol{w}_3\}$ 的坐标变换矩阵.

20. 设 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\},\mathcal{C}=\{\boldsymbol{c}_1,\boldsymbol{c}_2\}$ 和 $\mathcal{D}=\{\boldsymbol{d}_1,\boldsymbol{d}_2\}$ 是一个二维向量空间的基.
   a. 写出一个将 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P},\underset{\mathcal{D}\leftarrow\mathcal{C}}{P}$ 和 $\underset{\mathcal{D}\leftarrow\mathcal{B}}{P}$ 联系起来的方程，验证你的结果.
   b. [M] 针对 $\mathbb{R}^2$ 中的 3 个基（见 7~10 题），利用一个矩阵程序帮你找到这个方程或者检验你写出的方程.

练习题答案

1. 由于 P 的列是 $\mathcal{G}$－坐标向量，形如 $P\boldsymbol{x}$ 的向量一定是 $\mathcal{G}$－坐标向量，从而 P 满足方程（ii）.
2. 例 1 中求得的坐标向量表明

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}=[\,[\boldsymbol{b}_1]_{\mathcal{C}}\ \ [\boldsymbol{b}_2]_{\mathcal{C}}\,]=\begin{bmatrix}4&-6\\1&1\end{bmatrix}$$

从而

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}=\left(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}\right)^{-1}=\frac{1}{10}\begin{bmatrix}1&6\\-1&4\end{bmatrix}=\begin{bmatrix}0.1&0.6\\-0.1&0.4\end{bmatrix}$$

㊀ 习题 17 和习题 18 以及前几节中的 5 个相关习题的思想来自 Auburn 大学 Jack W.Rogers，Jr.写的一篇文章，此文章于 1995 年 8 月在国际线性代数协会的会议上宣读，见“Applications of Linear Algebra in Calculus”，*American Mathematical Monthly* 104(1), 1997.

4.8 差分方程中的应用

现在，功能强大的计算机被广泛地应用在各个领域. 越来越多的科学上和工程上的问题，在某种意义上讲，用离散的或数字化的数据来处理胜过用连续的数据来处理. 差分方程往往是分析这样的数据的合适工具，甚至当使用微分方程作连续过程的模型时，其数值解也常常由一个相关的差分方程得到.

本节中强调一些线性差分方程的基本性质，它们可通过线性代数得到很好的解释.

离散时间信号

离散时间信号的向量空间 $\mathbb{S}$ 在 4.1 节中引入，$\mathbb{S}$ 中一个**信号**是一个只定义在整数上的函数，同时可用一个数列将其直观化，即 $\{y_k\}$. 图 4-27 中展示出三个典型的信号，它们的通项分别是 $(0.7)^k, 1^k$ 和 $(-1)^k$.

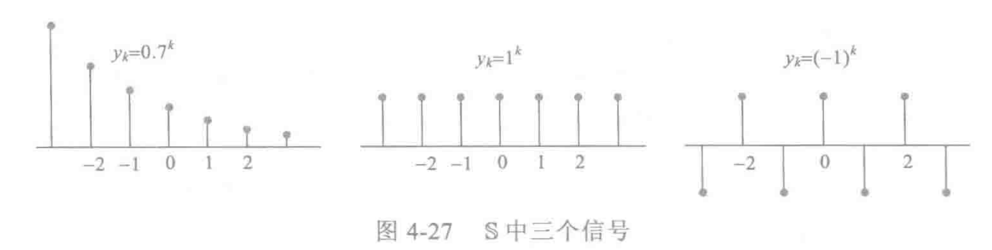


图 4-27　$\mathbb{S}$ 中三个信号

数字信号显然来自电学和控制系统工程学，但离散数据序列也来自生物学、物理学、经济学、人口统计学以及其他任何需要在离散时间区间测量或抽样的过程的领域. 如果一个过程从一个指定的时间开始，用形如 $(y_0, y_1, y_2, \cdots)$ 的序列去描述一个信号有时是方便的. 对于 $k<0$ 的 y_k 项，可以假设取值为 0 或者予以忽略.

例 1　光盘唱机中发出的清晰的声音是以每秒 44 100 次的速度从音乐中抽样而成的，见图 4-28. 在每次测量时，音乐信号的振幅用一个数字的形式记录下来，即 y_k. 最初的音乐是由各种频率的不同声音合成的，然而序列 $\{y_k\}$ 包含足够多的信息用来复制声音中的所有频率，最高达到大约每秒 20 000 个周期，这超出了人耳所能感觉到的范围.

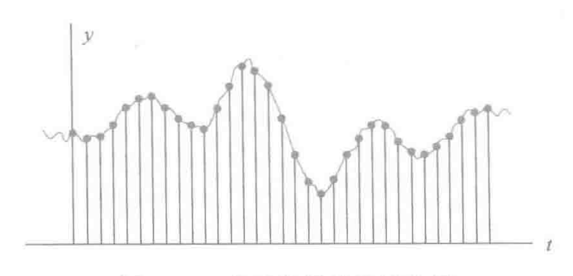


图 4-28　音乐信号的抽样数据 ■

信号空间 $\mathbb{S}$ 中的线性无关性

为了简化符号，我们考虑一个仅包含三个信号 $\{u_k\},\{v_k\}$ 和 $\{w_k\}$ 的集合 $\mathbb{S}$，当方程

$$c_1u_k + c_2v_k + c_3w_k = 0 \quad \text{对所有 } k \text{ 成立} \tag{1}$$

蕴涵 $c_1 = c_2 = c_3 = 0$ 时，$\{u_k\},\{v_k\},\{w_k\}$ 恰好是线性无关的. 这里说“对所有 k 成立”指对所有整数——正整数、负整数和 0 均成立. 我们可能考虑从 $k=0$ 开始的信号，例如，这时“对所有 k 成立”将表示对所有 $k \geqslant 0$ 的整数成立.

假设 c_1,c_2,c_3 满足（1）式，那么方程（1）对任意三个相邻的值 $k,k+1$ 和 $k+2$ 成立，这样（1）蕴涵

$$c_1u_{k+1} + c_2v_{k+1} + c_3w_{k+1} = 0 \quad \text{对所有 } k \text{ 成立}$$
$$c_1u_{k+2} + c_2v_{k+2} + c_3w_{k+2} = 0 \quad \text{对所有 } k \text{ 成立}$$

从而 c_1,c_2,c_3 满足

$$\begin{bmatrix} u_k & v_k & w_k \\ u_{k+1} & v_{k+1} & w_{k+1} \\ u_{k+2} & v_{k+2} & w_{k+2} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{对所有 } k \text{ 成立} \tag{2}$$

这个方程组的系矩阵称为信号的 **Casorati 矩阵**，这个矩阵的行列式称为 $\{u_k\},\{v_k\},\{w_k\}$ 的 **Casorati 行列式**. 如果对至少一个 k 值 Casorati 矩阵可逆，则（2）将蕴涵 $c_1 = c_2 = c_3 = 0$，这就证明这三个信号是线性无关的.

例 2　证明 $1^k,(-2)^k$ 和 3^k 是线性无关的信号. 见图 4-29.

解　Casorati 矩阵是

$$\begin{bmatrix} 1^k & (-2)^k & 3^k \\ 1^{k+1} & (-2)^{k+1} & 3^{k+1} \\ 1^{k+2} & (-2)^{k+2} & 3^{k+2} \end{bmatrix}$$

通过行变换可相当简单地证明这个矩阵总是可逆的，然而，若用 $k=0$ 代替 k，则会更快地行化简这个数值矩阵：

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 3 \\ 1 & 4 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & 2 \\ 0 & 3 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & 2 \\ 0 & 0 & 10 \end{bmatrix}$$

这个 Casorati 矩阵对 $k=0$ 可逆，所以 $1^k,(-2)^k$ 和 3^k 是线性无关的.

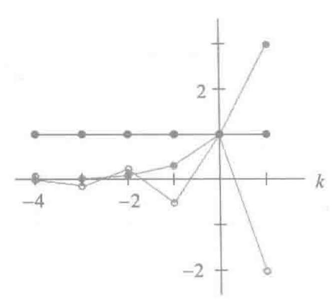


图 4-29　信号 1^k, $(-2)^k$ 和 3^k

若 Casorati 矩阵不可逆，相应的信号通过检测可能线性相关也可能不是线性相关（见习题33）. 但是可以证明，如果这些信号是同一个齐次差分方程（将在下面描述）的所有解，则 Casorati 矩阵对所有 k 是可逆的且这些信号是线性无关的，否则 Casorati 矩阵对所有 k 都不可逆且这些信号是线性相关的. 一个用线性变换方法的较好的证明可在学习指导中找到.

线性差分方程

给定数量 $a_0,\cdots,a_n$, a_0 和 a_n 不为零，给定一个信号 $\{z_k\}$，方程

$$a_0 y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = z_k \quad \text{对所有 } k \text{ 成立} \tag{3}$$

称为一个 n **阶线性差分方程**（或**线性递归关系**）. 为了简化，a_0 通常取为 1. 若 $\{z_k\}$ 是零序列，则方程是**齐次的**；否则，方程为**非齐次的**.

例 3 在数字信号处理中，像上面（3）那样的差分方程用来描述一个**线性滤波器**，$a_0,\cdots,a_n$ 称为滤波器系数，若将 $\{y_k\}$ 看作输入，$\{z_k\}$ 看作输出，则对应齐次方程的解是过滤掉的信号，并被变换成零信号，让我们向如下滤波器输入两个不同的信号：

$$0.35y_{k+2} + 0.5y_{k+1} + 0.35y_k = z_k$$

这里 0.35 是 $\sqrt{2}/4$ 的简写，第一个信号是由连续信号 $y = \cos(\pi t/4)$ 在整数值 t 抽样而生成，如图 4-30a 所示. 离散信号是 $\{y_k\} = \{\cdots,\cos(0),\cos(\pi/4),\cos(2\pi/4),\cos(3\pi/4),\cdots\}$.

为了简单，用 ±0.7 代替 $\pm\sqrt{2}/2$，从而

$$\{y_k\} = \{\cdots,\underset{\overset{\uparrow}{k=0}}{1},0.7,0,-0.7,-1,-0.7,0,0.7,1,0.7,0,\cdots\}$$

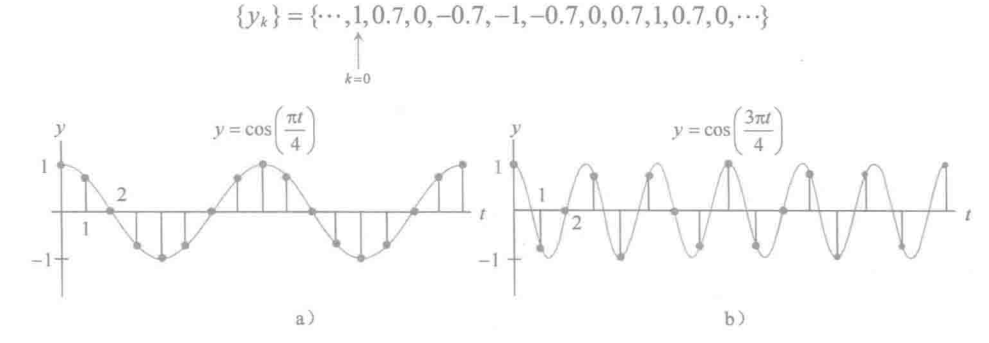


图 4-30 具有不同频率的离散信号

表 4-2 展示出输出序列 $\{z_k\}$ 的一个计算. 这里 0.35(0.7) 是 $(\sqrt{2}/4)(\sqrt{2}/2)=0.25$ 的简写，将 $\{z_k\}$ 移动一项，输出就成为 $\{y_k\}$.

表 4-2 计算一个滤波器的输出

k	y_k	y_{k+1}	y_{k+2}	$0.35y_k$	$+0.5y_{k+1}$	$+0.35y_{k+2}$	=	z_k
0	1	0.7	0	0.35(1)	+0.5(0.7)	+0.35(0)	=	0.7
1	0.7	0	−0.7	0.35(0.7)	+0.5(0)	+0.35(−0.7)	=	0
2	0	−0.7	−1	0.35(0)	+0.5(−0.7)	+0.35(−1)	=	−0.7
3	−0.7	−1	−0.7	0.35(−0.7)	+0.5(−1)	+0.35(−0.7)	=	−1
4	−1	−0.7	0	0.35(−1)	+0.5(−0.7)	+0.35(0)	=	−0.7
5	−0.7	0	0.7	0.35(−0.7)	+0.5(0)	+0.35(0.7)	=	0
⋮	⋮							⋮

于是一个不同的输入信号由更高的频率信号 $y=\cos(3\pi t/4)$ 生成，见图 4-30b. 用与前面相同的抽样率去抽样，我们得到一个新的输入序列：

$$\{w_k\}=\{\cdots,\underset{\substack{\uparrow\\ k=0}}{1},-0.7,0,0.7,-1,0.7,0,-0.7,1,-0.7,0\cdots\}$$

当将 $\{w_k\}$ 输入给滤波器，则输出是零序列，若一个滤波器使 $\{y_k\}$ 能够通过，但将高频的 $\{w_k\}$ 截掉，则称该滤波器为低通滤波器.

在许多应用中，序列 $\{z_k\}$ 由差分方程（3）的右端确定，满足（3）的一个 $\{y_k\}$ 称为这个方程组的一个**解**. 下一个例子表明对一个齐次方程如何求解. ■

例 4 齐次差分方程的解通常具有形式 $y_k=r^k$ 对某 r 成立，求下列方程的解.

$$y_{k+3}-2y_{k+2}-5y_{k+1}+6y_k=0 \quad 对所有 k 成立 \tag{4}$$

解 用 r^k 代替方程中的 y_k，并将左边分解因子.

$$r^{k+3}-2r^{k+2}-5r^{k+1}+6r^k=0 \tag{5}$$

$$r^k(r^3-2r^2-5r+6)=0$$

$$r^k(r-1)(r+2)(r-3)=0 \tag{6}$$

由于（5）等价于（6），r^k 满足差分方程（4）当且仅当 r 满足（6），于是 $1^k,(-2)^k$ 和 3^k 都是（4）的解，比如，为验证 3^k 是（4）的一个解，计算

$$3^{k+3}-2\cdot3^{k+2}-5\cdot3^{k+1}+6\cdot3^k=3^k(27-18-15+6)=0 \quad 对所有 k 成立$$ ■

一般而言，一个非零信号 r^k 满足齐次差分方程

$$y_{k+n}+a_1y_{k+n-1}+\cdots+a_{n-1}y_{k+1}+a_ny_k=0 \quad 对所有 k 成立$$

当且仅当 r 是**辅助方程**

$$r^n+a_1r^{n-1}+\cdots+a_{n-1}r+a_n=0$$

的一个根，我们将不考虑当 r 是辅助方程的重根的情形. 当这个辅助方程有复根，则差分方程具有形如 $s^k\cos k\omega$ 和 $s^k\sin k\omega$ 的解，其中 s 和 ω 是常数，这在例 3 中已出现过.

线性差分方程的解集

给定 $a_1,\cdots,a_n$，考虑映射 $T:\mathbb{S}\to\mathbb{S}$，将信号 $\{y_k\}$ 变换到信号 $\{w_k\}$，由下式给出

$$w_k=y_{k+n}+a_1y_{k+n-1}+\cdots+a_{n-1}y_{k+1}+a_ny_k$$

容易验证 T 是一个线性变换. 这蕴涵齐次方程

$$y_{k+n}+a_1y_{k+n-1}+\cdots+a_{n-1}y_{k+1}+a_ny_k=0 \quad 对所有 k 成立$$

的解集是 T 的核（经 T 映射到零信号的信号的集合），进而这个解集是 $\mathbb{S}$ 的一个子空间，任何解的线性组合仍然是解.

下一个定理是一个简单但基本的结论，它将引导出关于差分方程组解集的更多的信息.

定理 16 *若 $a_n\neq0$ 且 $\{z_k\}$ 给定，只要 $y_0,\cdots,y_{n-1}$ 给定，方程*

$$y_{k+n}+a_1y_{k+n-1}+\cdots+a_{n-1}y_{k+1}+a_ny_k=z_k \quad 对所有 k 成立 \tag{7}$$

有唯一解.

证 若 $y_0,\cdots,y_{n-1}$ 给定，利用（7）定义

$$y_n = z_0 - [a_1 y_{n-1} + \cdots + a_{n-1} y_1 + a_n y_0]$$

现在 $y_1,\cdots,y_n$ 明确了，再利用（7）定义 y_{n+1}．一般地，利用递归关系

$$y_{n+k} = z_k - [a_1 y_{k+n-1} + \cdots + a_n y_k] \tag{8}$$

对 $k \geqslant 0$ 定义 y_{n+k}．对 $k<0$，为了定义 y_k，利用递归关系

$$y_k = \frac{1}{a_n} z_k - \frac{1}{a_n}[y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1}] \tag{9}$$

这样生成了一个满足（7）式的信号．反之，对任何 k，满足（7）式的任何信号一定满足（8）和（9），所以（7）的解是唯一的． ■

定理 17 n 阶齐次线性差分方程

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = 0 \quad \text{对所有 } k \text{ 成立} \tag{10}$$

的解集 H 是一个 n 维向量空间．

证 我们早已解释过为什么 H 是 $\mathbb{S}$ 的一个子空间．对 H 中的 $\{y_k\}$，设 $F\{y_k\}$ 是 $\mathbb{R}^n$ 中的向量 $(y_0,y_1,\cdots,y_{n-1})$．容易证明 $F:H\to\mathbb{R}^n$ 是一个线性变换．任给 $\mathbb{R}^n$ 中向量 $(y_0,y_1,\cdots,y_{n-1})$，由定理 16 知存在 H 中唯一一个信号 $\{y_k\}$ 使得

$$F\{y_k\} = (y_0, y_1, \cdots, y_{n-1})$$

这说明 F 是由 H 到 $\mathbb{R}^n$ 上的一对一线性变换，即 F 是一个同构，从而 $\dim H = \dim \mathbb{R}^n = n$．（见 4.5 节习题 32）． ■

例 5 对差分方程

$$y_{k+3} - 2y_{k+2} - 5y_{k+1} + 6y_k = 0 \text{，对所有 } k \text{ 成立}$$

求其解集的一个基．

解 我们在线性代数中的工作现在真的要给我们回报了!从例 2 和例 4 知道 $1^k,(-2)^k$ 和 3^k 是线性无关解．一般地，直接证明一个信号的集合生成差分方程的解空间可能是困难的，但在这里因为有两个关键定理——定理 17 和 4.5 节中的基定理，所以使得我们的工作很容易完成，由定理 17 知，本例中方程的解空间恰好是三维的，再由 4.5 节中基定理知，n 维空间中含 n 个向量的线性无关集必然是一个基，所以 $1^k,(-2)^k$ 和 3^k 构成解空间的一个基． ■

描述（10）式的“通解”的标准方法是对所有解构成的子空间给出它的一个基，这样的基称为（10）的**基础解系**．实际上，如果我们能找到 n 个线性无关的信号满足（10），它们必然生成这个 n 维解空间，就像我们见到的上面的例题那样．

非齐次方程

非齐次差分方程

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = z_k \quad \text{对所有 } k \text{ 成立} \tag{11}$$

的通解能写成（11）的一个特解加上对应齐次差分方程（10）的一个基础解系的任意线性组合．这个结果类似于 1.5 节中关于 $A\boldsymbol{x}=\boldsymbol{b}$ 和 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集的关系，二者是类似的．这两个结果有相

同的意义：映射 $\boldsymbol{x} \mapsto A\boldsymbol{x}$ 是线性的，（11）中将信号 $\{y_k\}$ 变换成信号 $\{z_k\}$ 的映射也是线性的，见习题 35.

例 6 证明：信号 $y_k = k^2$ 满足差分方程

$$y_{k+2} - 4y_{k+1} + 3y_k = -4k \text{，对所有 } k \text{ 成立} \qquad (12)$$

然后给出这个方程所有解的一个刻画.

解 将（12）左端中的 y_k 用 k^2 代替.

$$\begin{aligned}(k+2)^2 - 4(k+1)^2 + 3k^2 & \\ &= (k^2+4k+4) - 4(k^2+2k+1) + 3k^2 \\ &= -4k\end{aligned}$$

所以 k^2 的确是（12）的一个解. 下一步是解齐次方程

$$y_{k+2} - 4y_{k+1} + 3y_k = 0 \qquad (13)$$

辅助方程为

$$r^2 - 4r + 3 = (r-1)(r-3) = 0$$

根为 $r = 1, 3$，所以齐次差分方程的两个解为 1^k 和 3^k，显然它们彼此不是倍数关系，所以它们是线性无关信号.（也可用 Casorati 检验）. 由定理 17，此解空间是 2 维的，所以 3^k 和 1^k 构成（13）的解集的一个基，该解集用非齐次差分方程（12）的特解形式写出来，得到（12）的通解：

$$k^2 + c_1 1^k + c_2 3^k \text{ 或 } k^2 + c_1 + c_2 3^k$$

图 4-31 给出两个解集的几何直观解释，图中每个点对应 $\mathbb{S}$ 中的一个信号.

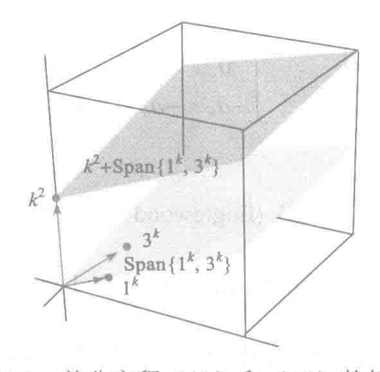


图 4-31 差分方程（12）和（13）的解集 ■

化简成一阶方程组

研究 n 阶齐次线性差分方程的现代方法是用等价的一阶差分方程组代替它，其中一阶差分方程写成如下形式：

$$\boldsymbol{x}_{k+1} = A\boldsymbol{x}_k \text{，对所有 } k \text{ 成立}$$

这里向量 $\boldsymbol{x}_k$ 在 $\mathbb{R}^n$ 中，A 是一个 $n \times n$ 矩阵.

这样的（向量值）差分方程的简单例子在 1.10 节中已经研究过. 进一步的例子将在 4.9 节和 5.6 节中给出.

例 7 将下列差分方程写成一个一阶方程组：

$$y_{k+3}-2y_{k+2}-5y_{k+1}+6y_k=0\text{，对所有 }k\text{ 成立}$$

解 对每个 k，设

$$\boldsymbol{x}_k=\begin{bmatrix}y_k\\y_{k+1}\\y_{k+2}\end{bmatrix}$$

由差分方程得 $y_{k+3}=-6y_k+5y_{k+1}+2y_{k+2}$，所以

$$\boldsymbol{x}_{k+1}=\begin{bmatrix}y_{k+1}\\y_{k+2}\\y_{k+3}\end{bmatrix}=\begin{bmatrix}0 & + & y_{k+1}+0\\0 & +0 & +\ y_{k+2}\\-6y_k & +5y_{k+1} & +2y_{k+2}\end{bmatrix}=\begin{bmatrix}0&1&0\\0&0&1\\-6&5&2\end{bmatrix}\begin{bmatrix}y_k\\y_{k+1}\\y_{k+2}\end{bmatrix}$$

即 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 对所有 k 成立，这里 $A=\begin{bmatrix}0&1&0\\0&0&1\\-6&5&2\end{bmatrix}$. ■

一般而言，方程

$$y_{k+n}+a_1y_{k+n-1}+\cdots+a_{n-1}y_{k+1}+a_ny_k=0\text{，对所有 }k\text{ 成立}$$

可重写成 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$，对所有 k 成立，这里

$$\boldsymbol{x}_k=\begin{bmatrix}y_k\\y_{k+1}\\\vdots\\y_{k+n-1}\end{bmatrix},\quad A=\begin{bmatrix}0&1&0&\cdots&0\\0&0&1&&0\\\vdots&&&\ddots&\vdots\\0&0&0&&1\\-a_n&-a_{n-1}&-a_{n-2}&\cdots&-a_1\end{bmatrix}$$

进一步阅读

Hamming, R. W., *Digital Filters*, 2nd ed. (Englewood Cliffs, NJ: Prentice-Hall, 1983), pp. 1–37.

Kelly, W. G., and A. C. Peterson, *Difference Equations*, 2nd ed. (San Diego: Harcourt-Academic Press, 2001).

Mickens, R. E., *Difference Equations*, 2nd ed. (New York: Van Nostrand Reinhold, 1990), pp. 88–141.

Oppenheim, A. V., and A. S. Willsky, *Signals and Systems*, 2nd ed. (Upper Saddle River, NJ: Prentice-Hall, 1997), pp. 1–14, 21–30, 38–43.

练习题

可以证明信号 $2^k, 3^k\sin\frac{k\pi}{2}$ 和 $3^k\cos\frac{k\pi}{2}$ 是

$$y_{k+3}-2y_{k+2}+9y_{k+1}-18y_k=0$$

的解，证明这些信号构成这个差分方程所有解集的一个基.

习题 4.8

验证习题 1 和习题 2 中的信号是相应差分方程的解.

1. $2^k, (-4)^k; y_{k+2} + 2y_{k+1} - 8y_k = 0$
2. $5^k, (-5)^k; y_{k+2} - 25y_k = 0$

证明习题 3~6 中的信号分别构成相应差分方程解集的一个基.

3. 习题 1 中的信号和方程.
4. 习题 2 中的信号和方程.
5. $(-2)^k, k(-2)^k; y_{k+2} + 4y_{k+1} + 4y_k = 0$
6. $4^k \cos\frac{k\pi}{2}, 4^k \sin\frac{k\pi}{2}; y_{k+2} + 16y_k = 0$

在习题 7~12 中，假设列出的信号是给出的差分方程的解，确定这些信号是否构成相应方程解空间的基.

7. $1^k, 2^k, (-2)^k; y_{k+3} - y_{k+2} - 4y_{k+1} + 4y_k = 0$
8. $(-1)^k, 2^k, 3^k; y_{k+3} - 4y_{k+2} + 1y_{k+1} + 6y_k = 0$
9. $2^k, 5^k \cos\frac{k\pi}{2}, 5^k \sin\frac{k\pi}{2}; y_{k+3} - 2y_{k+2} + 25y_{k+1} - 50y_k = 0$
10. $(-2)^k, k(-2)^k, 3^k; y_{k+3} + y_{k+2} - 8y_{k+1} - 12y_k = 0$
11. $(-1)^k, 2^k; y_{k+3} - 3y_{k+2} + 4y_k = 0$
12. $3^k, (-2)^k; y_{k+4} - 13y_{k+2} + 36y_k = 0$

在习题 13~16 中，分别求差分方程解空间的一个基.

13. $y_{k+2} - y_{k+1} + \frac{2}{9}y_k = 0$
14. $y_{k+2} - 5y_{k+1} + 6y_k = 0$
15. $6y_{k+2} + y_{k+1} - 2y_k = 0$
16. $y_{k+2} - 25y_k = 0$

在习题 17 和习题 18 中，涉及国民经济的一个简单模型，用下列差分方程描述

$$Y_{k+2} - a(1+b)Y_{k+1} + abY_k = 1 \qquad (14)$$

这里 Y_k 是第 k 年国民收入总和，a 为一个小于 1 的常数，称之为边际消费倾向，b 是一个正的调节常数，用来刻画消费性开支对一年的私人投资率的影响如何变化.[㊀]

17. 当 $a = 0.9, b = \frac{4}{9}$ 时，求（14）的通解，当 k 增加时，Y_k 如何变化？（提示：先求形如 $Y_k = T$ 的特解，这里 T 是一个常数，称为国民收入的平均水平.）
18. 当 $a = 0.9, b = 0.5$ 时，求（14）的通解.

一个轻的悬梁被间隔 10 英尺的 N 个支点支撑着，一个重 500 磅的法码挂在悬梁的一端，它距第一个支点 10 英尺，如图 4-32 所示. 设 y_k 表示第 k 个支点的弯矩，则 $y_1 = 5000$ 英尺-磅. 假设这个悬梁刚性地连接在第 N 个支点上并且此处弯矩为零，则各弯矩满足以下 3-弯矩方程：

$$y_{k+2} + 4y_{k+1} + y_k = 0 \text{ 对 } k = 1, 2, \cdots, N-2 \text{ 成立} \qquad (15)$$

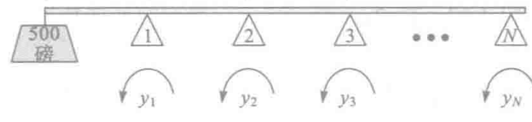


图 4-32 悬梁上的弯矩

19. 求差分方程（15）的通解. 验证你的答案.
20. 求满足边界条件 $y_1 = 5000, y_N = 0$ 的（15）的特解（答案涉及 N）.
21. 当一个信号通过在一个过程中（化学反应、通过管子的热流、活动机械手等）的一系列测量生成时，这个信号通常包括由测量误差造成的随机噪声. 预处理这些数据以便减少噪声的标准方法是将这个数据打磨，或叫过滤. 一个简单的过滤法是用两个相邻值的平均数代替每个 y_k 的移动平均数：

$$\frac{1}{3}y_{k+1} + \frac{1}{3}y_k + \frac{1}{3}y_{k-1} = z_k \text{ 对 } k = 1, 2, \cdots \text{均成立}$$

㊀ 例如，可参考 *Discrete Dynamical Systems*，James T. Sandefur (Oxford: Clarendon Press, 1990), pp.267-276，最原始的加速乘数模型由经济学家 P. A. Samuelson 提出.

假设对 $k=0,\cdots,14$, 一个信号 y_k 是:

$$9,5,7,3,2,4,6,5,7,6,8,10,9,5,7$$

利用过滤法计算 $z_1,\cdots,z_{13}$. 作一个原始信号和打磨过的信号的双重折线图.

22. 设 $\{y_k\}$ 是由连续信号 $2\cos\frac{\pi t}{4}+\cos\frac{3\pi t}{4}$ 在 $t=0,1,2,\cdots$ 处抽样生成的,如图 4-33 所示. 从 $k=0$ 开始 y_k 值为 3, 0.7, 0, –0.7, –3, –0.7, 0,0.7, 3, 0.7, 0, ⋯.这里 0.7 是 $\sqrt{2}/2$ 的简写.

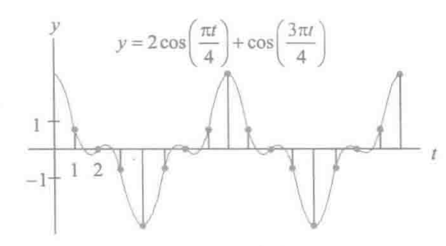


图 4-33 由 $2\cos\frac{\pi t}{4}+\cos\frac{3\pi t}{4}$ 得到的抽样数据

a. 当 $\{y_k\}$ 提供给例 3 中的滤波器时，计算输出信号 $\{z_k\}$.

b. 解释为什么（a）中输出与例 3 中的计算相关，有怎样的关系？

在习题 23 和 24 中，对适当的常数 a 和 b，涉及一个形如 $y_{k+1}-ay_k=b$ 的差分方程.

23. 10 000 美元的贷款每月有 1%的利息和 450 美元的月供. 一个月之后在 $k=1$ 时办理第一次付款. 对 $k=0,1,2,\cdots$, 设 y_k 是第 k 次月度付款刚办理后贷款的未付余额，则

$$y_1=10\,000\ +(0.01)10000-450$$

新余额 还贷额 附加利息 月供

a. 写出 $\{y_k\}$ 满足的差分方程.

b. [M]做一张表展示月份 k 时 k 与余额 y_k，列出你作这张表的程序和按键.

c. [M]当作完最后的付款时，k 为多少？最后一次的付款是多少？借款者共支付多少钱？

24. 在时间 $k=0$，办理了一个 1 000 美元的最初存款的储蓄账户. 这个储蓄账户每年付 6%的利息（每月利率为 0.005），这个利息按每月复利计算. 在最初存款之后，每月将 200 美元存到账户里. 对 $k=0,1,2\cdots$, 设 y_k 表示在时间 k 刚办理完存款后账户里的存款金额.

a. 写出 $\{y_k\}$ 满足的差分方程.

b. [M]做一张表，展示 k 与在月份 k 时账户里的总金额，k 取 $k=0$ 到 60. 列出你作这张表的程序和按键.

c. [M]两年（即 24 个月）之后，账户里有多少存款？4 年和 5 年之后呢？5 年的总利息是多少？

在习题 25~28 中，证明给出的信号是相应差分方程的解，然后求差分方程的通解.

25. $y_k=k^2; y_{k+2}+3y_{k+1}-4y_k=7+10k$

26. $y_k=1+k; y_{k+2}-6y_{k+1}+5y_k=-4$

27. $y_k=k-2; y_{k+2}-4y_k=8-3k$

28. $y_k=1+2k; y_{k+2}-25y_k=-48k-20$

将习题 29 和习题 30 中的差分方程用一阶方程组 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$（对任意 k 成立）的形式写出来.

29. $y_{k+4}+3y_{k+3}-8y_{k+2}+6y_{k+1}-2y_k=0$

30. $y_{k+3}-5y_{k+2}+8y_k=0$

31. 下列差分方程是 3 阶的吗？解释之.

$$y_{k+3}+5y_{k+2}+6y_{k+1}=0$$

32. 下列差分方程的阶是多少？解释你的答案.

$$y_{k+3}+a_1y_{k+2}+a_2y_{k+1}+a_3y_k=0$$

33. 设 $y_k=k^2, z_k=2k|k|$，信号 $\{y_k\}$ 和 $\{z_k\}$ 线性无关吗？对 $k=0, k=-1$ 和 $k=-2$，求相应的 Casorati 矩阵 $C(k)$，并讨论你的结果.

34. 设 f,g,h 是定义在全体实数上的线性无关函数,通过取整数上的函数值抽样构造 3 个信号:

$$u_k=f(k),\ v_k=g(k),\ w_k=h(k)$$

在 $\mathbb{S}$ 上这些信号一定是线性无关的吗？讨论之.

35. 设 a 和 b 是非零数，定义映射 $T\{y_k\}=\{w_k\}$，其中 $w_k=y_{k+2}+ay_{k+1}+by_k$，证明：$T$ 是一个由 $\mathbb{S}$

到 $\mathbb{S}$ 中的线性变换.

36. 设 V 是一个向量空间，$T:V\to V$ 是一个线性变换. 给定 $z\in V$，假设 $x_p\in V$ 满足 $T(x_p)=z$，设 u 是 T 的核中的任一向量，证明：$u+x_p$ 满足非齐次方程 $T(x)=z$.

37. 设 $\mathbb{S}_0$ 是所有形如 $(y_0,y_1,y_2,\ldots)$ 的序列的向量空间，定义从 $\mathbb{S}_0$ 到 $\mathbb{S}_0$ 的线性变换 T 和 D 如下：
$$T(y_0,y_1,y_2,\ldots)=(y_1,y_2,y_3,\ldots)$$
$$D(y_0,y_1,y_2,\ldots)=(0,y_0,y_1,y_2,\ldots)$$
证明 $TD=I$（I 是 $\mathbb{S}_0$ 上的恒等变换）但 $DT\neq I$.

练习题答案

检查 Casorati 矩阵：

$$C(k)=\begin{bmatrix}2^k & 3^k\sin\frac{k\pi}{2} & 3^k\cos\frac{k\pi}{2}\\ 2^{k+1} & 3^{k+1}\sin\frac{(k+1)\pi}{2} & 3^{k+1}\cos\frac{(k+1)\pi}{2}\\ 2^{k+2} & 3^{k+2}\sin\frac{(k+2)\pi}{2} & 3^{k+2}\cos\frac{(k+2)\pi}{2}\end{bmatrix}$$

设 $k=0$，将矩阵进行行化简，证明它有 3 个主元位置，从而矩阵可逆：

$$C(0)=\begin{bmatrix}1&0&1\\2&3&0\\4&0&-9\end{bmatrix}\sim\begin{bmatrix}1&0&1\\0&3&-2\\0&0&-13\end{bmatrix}$$

当 $k=0$ 时，Casorati 矩阵是可逆的，所以这些信号是线性无关的，因为有 3 个信号，差分方程的解空间 H 是 3 维的（定理 17），由基定理知，这 3 个信号构成 H 的一个基.

4.9 马尔可夫链中的应用

本节中描述的马尔可夫链，在许多学科如生物学、商业、化学、工程学及物理学等领域中被用来做数学模型. 在每种情形中，该模型习惯上用来描述用同一种方法进行多次的实验或测量，实验中每次测试的结果属于几个指定的可能结果之一，每次测试结果仅依赖于最接近的前一次测试.

例如，若每年要统计一个城市及其郊区的人口，像

$$x_0=\begin{bmatrix}0.60\\0.40\end{bmatrix}\tag{1}$$

这样的向量可以显示 60%的人口住在这个城市中，40%的人口住在郊区. x_0 中的小数加起来等于 1 是因为它们说明这个地区的总人口，在此对我们的目的而言，用百分数表示比用人口总数表示更方便.

一个具有非负分量且各分量的数值相加等于 1 的向量称为**概率向量**；**随机矩阵**是各列向量均为概率向量的方阵；**马尔可夫链**是一个概率向量序列 $x_0,x_1,x_2,\cdots$ 和一个随机矩阵 P，使得

$$x_1=Px_0,\ x_2=Px_1,\ x_3=Px_2,\cdots$$

于是马尔可夫链可用一阶差分方程来刻画：

$$x_{k+1}=Px_k\ ,k=0,1,2,\cdots$$

当向量在 $\mathbb{R}^n$ 中的一个马尔可夫链描述一个系统或实验的序列时，$\boldsymbol{x}_k$ 中的数值分别列出系统在 n 个可能状态中的概率，或实验结果是 n 个可能结果之一的概率. 因此，$\boldsymbol{x}_k$ 通常称为**状态向量**.

例 1 在 1.10 节中，我们研究过一个人口在城市与郊区之间移动的模型，见图 4-34，在大城市地区的这两个部分之间，每年的移民由移民矩阵 M 控制：

$$\begin{array}{ccc} \text{由} & \text{城市} & \text{郊区} & \text{去} \end{array}$$
$$M=\begin{bmatrix}0.95 & 0.03\\ 0.05 & 0.97\end{bmatrix}\begin{array}{l}\text{城市}\\ \text{郊区}\end{array}$$

即每年有 5%的城市人口流动到郊区，有 3%的郊区人口流动到城市，M 的列是概率向量，所以 M 是一个随机矩阵. 假设在 2000 年这个地区的城市人口为 600 000，郊区人口为 400 000，所以这个地区原来的人口分布由上面（1）中的 $\boldsymbol{x}_0$ 给出，2001 年人口的分布是什么？2002 年呢？

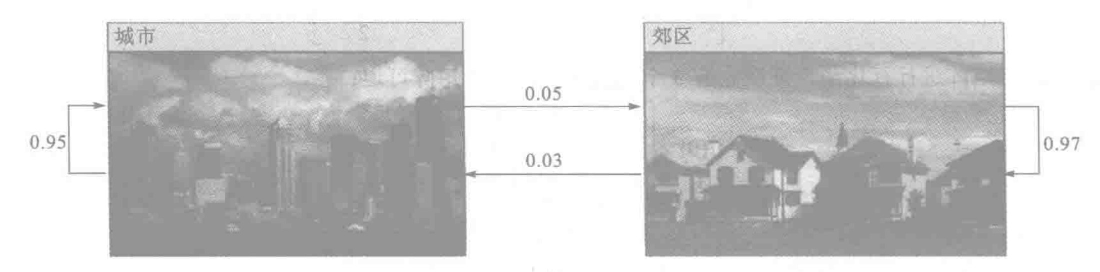


图 4-34 每年城市与郊区之间移民的百分比

解 在 1.10 节的例 3 中，我们看到 1 年之后，人口向量 $\begin{bmatrix}600\,000\\400\,000\end{bmatrix}$ 变为

$$\begin{bmatrix}0.95 & 0.03\\ 0.05 & 0.97\end{bmatrix}\begin{bmatrix}600\,000\\ 400\,000\end{bmatrix}=\begin{bmatrix}582\,000\\ 418\,000\end{bmatrix}$$

如果用总人口 100 万除以方程两边，再利用事实 $kM\boldsymbol{x}=M(k\boldsymbol{x})$，得

$$\begin{bmatrix}0.95 & 0.03\\ 0.05 & 0.97\end{bmatrix}\begin{bmatrix}0.600\\ 0.400\end{bmatrix}=\begin{bmatrix}0.582\\ 0.418\end{bmatrix}$$

向量 $\boldsymbol{x}_1=\begin{bmatrix}0.582\\0.418\end{bmatrix}$ 给出 2001 年的人口分布，即该地区 58.2%的人口住在城市，41.8%的人口住在郊区，类似地，2002 年的人口分布由 $\boldsymbol{x}_2$ 给出，这里

$$\boldsymbol{x}_2=M\boldsymbol{x}_1=\begin{bmatrix}0.95 & 0.03\\ 0.05 & 0.97\end{bmatrix}\begin{bmatrix}0.582\\ 0.418\end{bmatrix}=\begin{bmatrix}0.565\\ 0.435\end{bmatrix}$$ ■

例 2 假设在某一固定选区国会选举的投票结果用 $\mathbb{R}^3$ 中的向量 $\boldsymbol{x}$ 表示为

$$\boldsymbol{x}=\begin{bmatrix}\text{民主党得票率}(D)\\ \text{共和党得票率}(R)\\ \text{自由党得票率}(L)\end{bmatrix}$$

假设我们用这种类型的向量每两年记录一次国会选举的结果，同时每次选举的结果仅依赖前一次选举的结果，于是刻画每两年选举的向量构成的序列是一个马尔可夫链，对此链，作为一个随机矩阵的例子，取

$$\begin{array}{c} \text{从} \quad D \quad\quad R \quad\quad L \quad \text{到} \\ P=\begin{bmatrix} 0.70 & 0.10 & 0.30 \\ 0.20 & 0.80 & 0.30 \\ 0.10 & 0.10 & 0.40 \end{bmatrix} \begin{matrix} D \\ R \\ L \end{matrix} \end{array}$$

标志为“D”的第一列中的数值刻画在一次选举中为民主党投票的人在下一次选举中将如何投票的百分比．这里我们已经假设 70%的人在下一次选举中再一次投“D”的票，20%的人将投“R”的票，10%的人将投“L”的票，对 P 的其他两列有类似的解释，图 4-35 给出这个矩阵的一个图表．

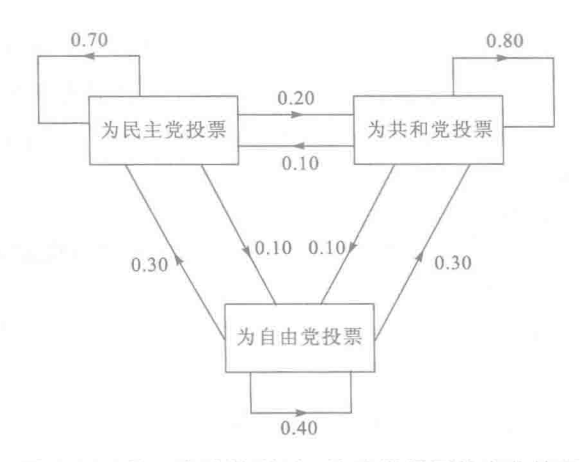


图 4-35　从一次选举到下一次选举投票的变化情况

如果这些“转换”百分比从一次选举到下一次选举多年保持为常数，则那些给出投票结果的向量的序列构成一个马尔可夫链．假设在一次选举中，结果为

$$\boldsymbol{x}_0=\begin{bmatrix} 0.55 \\ 0.40 \\ 0.05 \end{bmatrix}$$

确定下一次可能的结果和再下一次可能的结果．

解　下一次选举的结果由状态向量 $\boldsymbol{x}_1$ 描述，再下一次选举的结果由 $\boldsymbol{x}_2$ 描述，这里

$$\boldsymbol{x}_1=P\boldsymbol{x}_0=\begin{bmatrix} 0.70 & 0.10 & 0.30 \\ 0.20 & 0.80 & 0.30 \\ 0.10 & 0.10 & 0.40 \end{bmatrix}\begin{bmatrix} 0.55 \\ 0.40 \\ 0.05 \end{bmatrix}=\begin{bmatrix} 0.440 \\ 0.445 \\ 0.115 \end{bmatrix} \begin{matrix} \text{44\%将投 } D \text{ 的票} \\ \text{44.5\%将投 } R \text{ 的票} \\ \text{11.5\%将投 } L \text{ 的票} \end{matrix}$$

$$x_2 = Px_1 = \begin{bmatrix} 0.70 & 0.10 & 0.30 \\ 0.20 & 0.80 & 0.30 \\ 0.10 & 0.10 & 0.40 \end{bmatrix} \begin{bmatrix} 0.440 \\ 0.445 \\ 0.115 \end{bmatrix} = \begin{bmatrix} 0.3870 \\ 0.4785 \\ 0.1345 \end{bmatrix} \begin{matrix} 38.7\%\text{将投}D\text{的票} \\ 47.8\%\text{将投}R\text{的票} \\ 13.5\%\text{将投}L\text{的票} \end{matrix}$$

为了搞清楚为什么 x_1 事实上给出了下一次选举的结果，假设 1000 个人在“第一次”选举中投票，550 人投 D 的票，400 人投 R 的票，50 人投 L 的票（见 x_0 中的百分比）. 在下一次选举中，550 人中的 70%将再一次投 D 的票，400 人中的 10%将从 R 转投 D，50 人中的 30%将从 L 转投 D，于是 D 的总得票数为

$$0.70(550) + 0.10(400) + 0.30(50) = 385 + 40 + 15 = 440 \tag{2}$$

于是下一次 D 的候选人将得 44%的选票，(2)中的计算本质上与计算 x_1 中第一个数值是相同的，对 x_1 中其他数值以及 x_2 中的数值等可以作类似的计算. ■

预言遥远的未来

马尔可夫链最有趣的方面是对该链长期行为的研究. 例如，在例 2 中经过多次选举以后，关于投票的情况我们能说些什么？（假设从一次选举到下一次选举，给定的随机矩阵连续描述转换百分比.）另外，从长远看，例 1 中的人口分布将有什么样的结果？在回答这些问题之前，我们先讨论一个数值的例子.

例 3 令 $P = \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.8 & 0.3 \\ 0.2 & 0 & 0.4 \end{bmatrix}, x_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$，考虑一个系统，它的状态由马尔可夫链 $x_{k+1} = Px_k$（$k = 0,1,\cdots$）描述. 随着时间的流逝，这个系统将有什么结果？为此，计算状态向量 $x_1, \cdots x_{15}$ 来寻找结果.

解 后面向量中的数值保留 4 位或 5 位有效数字.

$$x_1 = Px_0 = \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.8 & 0.3 \\ 0.2 & 0 & 0.4 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.3 \\ 0.2 \end{bmatrix}$$

$$x_2 = Px_1 = \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.8 & 0.3 \\ 0.2 & 0 & 0.4 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.3 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 0.37 \\ 0.45 \\ 0.18 \end{bmatrix}$$

$$x_3 = Px_2 = \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.8 & 0.3 \\ 0.2 & 0 & 0.4 \end{bmatrix} \begin{bmatrix} 0.37 \\ 0.45 \\ 0.18 \end{bmatrix} = \begin{bmatrix} 0.329 \\ 0.525 \\ 0.146 \end{bmatrix}$$

继续用这种方法，有

$$x_4 = \begin{bmatrix} 0.3133 \\ 0.5625 \\ 0.1242 \end{bmatrix}, \quad x_5 = \begin{bmatrix} 0.3064 \\ 0.5813 \\ 0.1123 \end{bmatrix}, \quad x_6 = \begin{bmatrix} 0.3032 \\ 0.5906 \\ 0.1062 \end{bmatrix}, \quad x_7 = \begin{bmatrix} 0.3016 \\ 0.5953 \\ 0.1031 \end{bmatrix}$$

$$x_8=\begin{bmatrix}0.3008\\0.5977\\0.1016\end{bmatrix},\ x_9=\begin{bmatrix}0.3004\\0.5988\\0.1008\end{bmatrix},\ x_{10}=\begin{bmatrix}0.3002\\0.5994\\0.1004\end{bmatrix},\ x_{11}=\begin{bmatrix}0.3001\\0.5997\\0.1002\end{bmatrix}$$

$$x_{12}=\begin{bmatrix}0.30005\\0.59985\\0.10010\end{bmatrix},\ x_{13}=\begin{bmatrix}0.30002\\0.59993\\0.10005\end{bmatrix},\ x_{14}=\begin{bmatrix}0.30001\\0.59996\\0.10002\end{bmatrix},\ x_{15}=\begin{bmatrix}0.30001\\0.59998\\0.10001\end{bmatrix}$$

这些向量似乎是逼近 $q=\begin{bmatrix}0.3\\0.6\\0.1\end{bmatrix}$ 的. 这些概率由 k 的一个值到下一个值几乎不改变，注意到下列计算是精确的（没有舍入误差）：

$$Pq=\begin{bmatrix}0.5&0.2&0.3\\0.3&0.8&0.3\\0.2&0&0.4\end{bmatrix}\begin{bmatrix}0.3\\0.6\\0.1\end{bmatrix}=\begin{bmatrix}0.15+0.12+0.03\\0.09+0.48+0.03\\0.06+0\quad+0.04\end{bmatrix}=\begin{bmatrix}0.30\\0.60\\0.10\end{bmatrix}=q$$

若系统处于状态 q，则从一次测量到下一次测量，系统没有变化. ■

稳态向量

若 P 是一个随机矩阵，则相对于 P 的**稳态向量**（或**平衡向量**）是一个满足

$$Pq=q$$

的概率向量 q. 可以证明每一个随机矩阵有一个稳态向量，上面的例 3 中，q 是 P 的一个稳态向量.

例 4 在例 1 中，概率向量 $q=\begin{bmatrix}0.375\\0.625\end{bmatrix}$ 是人口迁移矩阵 M 的一个稳态向量. 这是因为

$$Mq=\begin{bmatrix}0.95&0.03\\0.05&0.97\end{bmatrix}\cdot\begin{bmatrix}0.375\\0.625\end{bmatrix}=\begin{bmatrix}0.35625+0.01875\\0.01875+0.60625\end{bmatrix}=\begin{bmatrix}0.375\\0.625\end{bmatrix}=q$$ ■

在例 1 中，若大城市地区的总人口是 100 万，则由例 4，q 将对应有 375 000 人在城市，有 625 000 人在郊区，在一年的年底，从城市迁出的人口是

$$(0.05)\,(375\,000)=18\,750$$

人，从郊区迁进城市的人口是 $(0.03)\,(625\,000)=18\,750$ 人. 结果是，城市里的人口保持不变，类似地，郊区里的人口也是稳定的.

下一个例子说明如何求稳态向量.

例 5 令 $P=\begin{bmatrix}0.6&0.3\\0.4&0.7\end{bmatrix}$，求 P 的稳态向量.

解 首先，解方程 $Px=x$.

$$Px-x=\mathbf{0}$$

$$Px-Ix=\mathbf{0}\quad \text{回顾 1.4 节有 } Ix=x$$

$$(P-I)x=\mathbf{0}$$

对上面的 P，

$$P-I=\begin{bmatrix}0.6 & 0.3\\0.4 & 0.7\end{bmatrix}-\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}=\begin{bmatrix}-0.4 & 0.3\\0.4 & -0.3\end{bmatrix}$$

为求 $(P-I)\boldsymbol{x}=\boldsymbol{0}$ 的所有解，将增广矩阵作行化简.

$$\begin{bmatrix}-0.4 & 0.3 & 0\\0.4 & -0.3 & 0\end{bmatrix}\sim\begin{bmatrix}-0.4 & 0.3 & 0\\0 & 0 & 0\end{bmatrix}\sim\begin{bmatrix}1 & -3/4 & 0\\0 & 0 & 0\end{bmatrix}$$

于是 $x_1=\dfrac{3}{4}x_2, x_2$ 为自由变量，通解为 $x_2\begin{bmatrix}3/4\\1\end{bmatrix}$.

其次，对此解空间选一个简单的基. 一个显然的选择是 $\begin{bmatrix}3/4\\1\end{bmatrix}$，但一个更好的没有分数的选择是 $\boldsymbol{w}=\begin{bmatrix}3\\4\end{bmatrix}$（对于 $x_2=4$）.

最后，在 $P\boldsymbol{x}=\boldsymbol{x}$ 的全体解的集合中求一个概率向量，这是简单的，因为每个解均为上面 w 的一个倍数，将 $\boldsymbol{w}$ 除以其数值之和得

$$\boldsymbol{q}=\begin{bmatrix}3/7\\4/7\end{bmatrix}$$

检验之，计算

$$P\boldsymbol{q}=\begin{bmatrix}6/10 & 3/10\\4/10 & 7/10\end{bmatrix}\begin{bmatrix}3/7\\4/7\end{bmatrix}=\begin{bmatrix}18/70+12/70\\12/70+28/70\end{bmatrix}=\begin{bmatrix}30/70\\40/70\end{bmatrix}=\boldsymbol{q}$$

■

下一个定理将证明，例 3 中产生的结果是许多随机矩阵的一个典型代表. 我们说一个随机矩阵是**正则**的，如果矩阵的某次幂 P^k 仅包含正的数值，对例 3 中的 P，有

$$P^2=\begin{bmatrix}0.37 & 0.26 & 0.33\\0.45 & 0.70 & 0.45\\0.18 & 0.04 & 0.22\end{bmatrix}$$

由于 P^2 中每个数是严格正的，故 P 是一个正规随机矩阵.

另外，我们说一个向量序列 $\{\boldsymbol{x}_k: k=1, 2, \cdots\}$ 当 $k\to\infty$ 时，收敛到一个向量 $\boldsymbol{q}$，如果当 k 充分大时，$\boldsymbol{x}_k$ 中的数值无限接近 $\boldsymbol{q}$ 中对应的数值.

定理 18 若 P 是一个 $n\times n$ 的正则随机矩阵，则 P 具有唯一的稳态向量 $\boldsymbol{q}$. 进一步，若 $\boldsymbol{x}_0$ 是任一个起始状态，且 $\boldsymbol{x}_{k+1}=P\boldsymbol{x}_k$, $k=0,1,2,\cdots$, 则当 $k\to\infty$ 时，马尔可夫链 $\{\boldsymbol{x}_k\}$ 收敛到 $\boldsymbol{q}$.

这个定理的证明可在马尔可夫链方面的标准教科书中找到，这个定理的奇妙之处在于初始状态对马尔可夫链的长期行为没有影响. 稍后（在 5.2 节）你将看到为什么这种情况对这里研究的多个随机矩阵是真实的.

例 6 在例 2 中，假设选举结果构成一个马尔可夫链. 问从现在开始经过多年若干次的选举之后，投票者可能为共和党候选人投票的百分比是多少?

解 若用手工计算，错误方法是选某初始向量 $\boldsymbol{x}_0$，再对充分大的 k 计算 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k$，这样没办法知道要计算多少向量，并且你不能把握 $\boldsymbol{x}_k$ 中数值的极限值.

正确方法是计算稳态向量再借助定理 18，给定例 2 中的矩阵 P，通过对角线上每个数值减去 1 得到 $P-I$，再将增广矩阵进行行化简：

$$[(P-I)\ \ \mathbf{0}]=\begin{bmatrix}-0.3 & 0.1 & 0.3 & 0\\ 0.2 & -0.2 & 0.3 & 0\\ 0.1 & 0.1 & -0.6 & 0\end{bmatrix}$$

回顾前面的工作，通过每一行乘以 10，可以使运算简化.[㊀]

$$\begin{bmatrix}-3 & 1 & 3 & 0\\ 2 & -2 & 3 & 0\\ 1 & 1 & -6 & 0\end{bmatrix}\sim\begin{bmatrix}1 & 0 & -9/4 & 0\\ 0 & 1 & -15/4 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}$$

$(P-I)\boldsymbol{x}=\mathbf{0}$ 的通解为 $x_1=\dfrac{9}{4}x_3$,$x_2=\dfrac{15}{4}x_3$,x_3 是自由变量. 选 $x_3=4$，得到解空间的一组基，它的每个数值是整数，由此容易求得隐态向量，它的数值之和为 1：

$$\boldsymbol{w}=\begin{bmatrix}9\\ 15\\ 4\end{bmatrix},\ \boldsymbol{q}=\begin{bmatrix}9/28\\ 15/28\\ 4/28\end{bmatrix}\approx\begin{bmatrix}0.32\\ 0.54\\ 0.14\end{bmatrix}$$

$\boldsymbol{q}$ 中的元素刻画由现在开始多年之后进行的一次选举中得票数的分布（假设这个随机矩阵连续描述从一次选举到下一次选举的变化情况）. 这样，最终大约 54%的选票被共和党候选人得到.

■

数值计算的注解 你可能已注意到若 $\boldsymbol{x}_{k+1}=P\boldsymbol{x}_k$, $k=0,1,2,\cdots$, 则

$$\boldsymbol{x}_2=P\boldsymbol{x}_1=P(P\boldsymbol{x}_0)=P^2\boldsymbol{x}_0$$

一般地，

$$\boldsymbol{x}_k=P^k\boldsymbol{x}_0,\ k=0,1,2,\cdots$$

为了计算一个向量，比如 $\boldsymbol{x}_3$，则将 $\boldsymbol{x}_1,\boldsymbol{x}_2$ 先算出来再计算 $\boldsymbol{x}_3$ 也只需要很少的算术运算，而先计算 P^3 和 $P^3\boldsymbol{x}_0$ 则不然. 然而，若 P 很小——比如是 30×30 的矩阵，对这两种方法，机器计算 $\boldsymbol{x}_3$ 的时间没有什么区别，计算 $P^3\boldsymbol{x}_0$ 的指令可以作为首选，因为它需要较少的键击次数.

练习题

1. 假设一个城市地区的居民按照例 1 中移民矩阵的概率进行迁移，假设某个居民“随机”选择去向. 则一确定年度的状态向量可以理解为给出了当时这个人是城市居民还是郊区居民的概率.

   a. 假设现在这个人是城市居民，所以 $\boldsymbol{x}_0=\begin{bmatrix}1\\ 0\end{bmatrix}$，这个人下一年住在郊区的可能性有多大？

   b. 这个人两年后住在郊区的可能性有多大？

㊀ 注：不仅仅对 P 乘以 10，事实上，是对方程 $(P-I)\boldsymbol{x}=\mathbf{0}$ 的增广矩阵乘以 10.

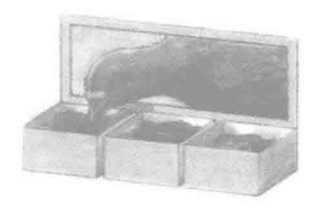
2. 令 $P=\begin{bmatrix}0.6 & 0.2\\0.4 & 0.8\end{bmatrix}$, $\boldsymbol{q}=\begin{bmatrix}0.3\\0.7\end{bmatrix}$，向量 $\boldsymbol{q}$ 是否为 P 的稳态向量？

3. 例 1 中多年之后住在郊区的人口的百分比是多少？

习题 4.9

1. 一个偏僻的小村子从两个广播站接收无线电广播，其中一个是新闻广播站，另一个是音乐广播站. 在广播站每半小时一次的休息后，接收新闻广播的听众中，有 70%的人还将保持收听新闻，有 30%的人在广播站休息时换成接收音乐广播. 接收音乐广播的听众中，当广播站休息时，有 60%的人将换成收听新闻广播. 有 40%的人还将保持听音乐. 假设在上午 8∶15，每个人都正在收听新闻.
   a. 给出描述无线电听众在每个广播站休息时趋向变换广播站的随机矩阵，将行、列做标记.
   b. 给出初始状态向量.
   c. 听众在上午 9∶25 收听音乐广播的百分率是多少？（在上午 8∶30 和 9∶00 广播站休息之后.）

2. 一个实验室动物每天可以吃三种食物中的任一种，实验室记录表明，在一次测试中，如果这个动物选择一种食物，在下一次测试中它选择同样食物的概率是 60%，下一次测试中它选择其他两种食物的概率均为 20%.
   a. 对此情形，随机矩阵是怎样的？
   b. 若这个动物在一次初始测试中选 1 号食物，在初始测试后的第二次测试中，它选 2 号食物的概率是多少？



3. 在任意给定的一天，一个学生要么健康，要么生病. 在所有今天健康的学生中，有 95%的人明天仍然健康，在所有今天在生病的学生中，有 55%的人明天仍然生病.
   a. 对此情形，随机矩阵是什么？
   b. 假设在星期一有 20%的学生是生病的，在星期二，学生中可能生病人数的百分比是多少？星期三呢？
   c. 如果一个学生今天是健康的，问该学生从现在开始两天后仍然健康的概率是多大？

4. 对任意给定的一天，哥伦布地区的天气要么是好，要么是中等，要么是差. 若今天天气好，则明天的天气有 40%的可能是好的，有 30%的可能是中等的，有 30%的可能是差的. 若今天的天气是中等的，则明天的天气有 50%的可能是好的，有 20%的可能是中等的. 最后，若今天天气差，则明天的天气有 30%的可能是好的，有 40%的可能是中等的.
   a. 对此情形，随机矩阵是怎样的？
   b. 假设今天是好天气的可能性为 50%，是中等天气的可能性为 50%，明天是差天气的可能性是多少？
   c. 假设对星期一的天气预报是 60%为中等天气，40%为差天气，在星期三是好天气的可能性是多少？

在习题 5~8 中，求稳态向量.

5. $\begin{bmatrix}0.1 & 0.5\\0.9 & 0.5\end{bmatrix}$

6. $\begin{bmatrix}0.4 & 0.8\\0.6 & 0.2\end{bmatrix}$

7. $\begin{bmatrix}0.7 & 0.1 & 0.1\\0.2 & 0.8 & 0.2\\0.1 & 0.1 & 0.7\end{bmatrix}$

8. $\begin{bmatrix}0.4 & 0.5 & 0.8\\0 & 0.5 & 0.1\\0.6 & 0 & 0.1\end{bmatrix}$

9. 确定 $P=\begin{bmatrix}0.2 & 1\\0.8 & 0\end{bmatrix}$ 是否是一个正则的随机矩阵？

10. 确定 $P=\begin{bmatrix}1 & 0.3\\0 & 0.7\end{bmatrix}$ 是否是一个正则的随机矩阵？

11. a. 对习题 1 中的马尔可夫链，求其稳态向量.
    b. 经过很长时间之后，听众中收听新闻的比例是多少？

12. 参考习题 2，多次测试后，这个动物更喜欢哪一种食物？

13. a. 对习题 3 中的马尔可夫链，求其稳态向量.
    b. 对某个确定的人，经过许多天之后，生病的概率是多少？若今天这个人正在发病，对这个结果有影响吗？

14. 参考习题 4，从长远看，对任给的一天，哥伦布地区的天气是好天气的可能性有多大？

15. [M]加州财政部的人口研究部门为下列移民矩阵提供了数据，这个矩阵刻画在 1989 年间美国人口的流动情况，在 1989 年，大约总人口的 11.7%居住在加州. 如果给出的移民概率多年保持不变，问总人口中最终居住在加州的百分比是多少？

由 加州	美国其他地区	到
0.982 1	0.002 9	加州
0.017 9	0.997 1	美国其他地区

16. [M]在底特律，Hertz 出租车公司大约有 2000 辆小汽车，租出和返回位置的模式由下表中分数给出. 在一天中，由市中心位置大约有多少辆小汽车将被租出或准备租出？汽车租出：

由 City 机场	市中心	Metro 机场	回到
0.90	0.01	0.09	City 机场
0.01	0.90	0.01	市中心
0.09	0.09	0.90	Metro 机场

17. 令 P 是一个 $n\times n$ 随机矩阵，下列论证表明方程 $P\boldsymbol{x}=\boldsymbol{x}$ 有一个非平凡解.（事实上，存在一个具有非负数值的稳态解，其证明在一些高等教材中给出.）证明下列每个论断.（注意用一个定理.）
    a. 若 $P-I$ 的其余行均加到最下一行，则结果为零行.
    b. $P-I$ 的行线性相关.
    c. $P-I$ 的行空间的维数小于 n.
    d. $P-I$ 有一个非平凡的零空间.

18. 证明：每一个 2×2 随机矩阵至少有一个稳态向量. 任意这样的矩阵均能写成形如 $P=\begin{bmatrix}1-\alpha & \beta\\ \alpha & 1-\beta\end{bmatrix}$ 的形式，这里 α 与 β 是 0 到 1 之间的常数.（若 $\alpha=\beta=0$，则存在两个线性无关的稳态向量，否则只有一个稳态向量.）

19. 令 S 是一个 $1\times n$ 行矩阵，每列中只有 1.

$$S=\begin{bmatrix}1 & 1 & \cdots 1\end{bmatrix}$$

    a. 解释为什么 $\mathbb{R}^n$ 中的向量是概率向量当且仅当它的数值非负且 $S\boldsymbol{x}=1$.（像 $S\boldsymbol{x}$ 这样的 1×1 矩阵通常不写矩阵括号符号.）
    b. 令 P 一个 $n\times n$ 随机矩阵，解释为什么 $SP=S$.
    c. 令 P 是一个 $n\times n$ 随机矩阵，$\boldsymbol{x}$ 是一个概率向量，证明：$P\boldsymbol{x}$ 也是一个概率向量.

20. 利用习题 19 证明：若 P 是一个 $n\times n$ 随机矩阵，则 P^2 也是.

21. [M]检验正则随机矩阵的幂.
    a. 计算 $P^k, k=2,3,4,5$.

$$P=\begin{bmatrix}0.335\,5 & 0.368\,2 & 0.306\,7 & 0.038\,9\\ 0.266\,3 & 0.272\,3 & 0.327\,7 & 0.545\,1\\ 0.193\,5 & 0.150\,2 & 0.158\,9 & 0.239\,5\\ 0.204\,7 & 0.209\,3 & 0.206\,7 & 0.176\,5\end{bmatrix}$$

    计算结果保留 4 位小数，当 k 增加时，P^k 的列将有什么结果？计算 P 的稳态向量.
    b. 计算 $Q^k, k=10,20,\cdots,80$.

$$Q=\begin{bmatrix}0.97 & 0.05 & 0.10\\ 0 & 0.90 & 0.05\\ 0.03 & 0.05 & 0.85\end{bmatrix}$$

    （使 Q^k 的稳定性达到 4 位小数可能需要 $k=116$ 或更大.）计算 Q 的稳态向量. 对任意正则的随机矩阵，猜想什么结果可能是正确的.
    c. 利用定理 18 解释在（a）和（b）中你所发现的结果.

22. [M]对比求正则随机矩阵 P 的稳态向量 $\boldsymbol{q}$ 的两种方法：
   （1）像例 5 中那样计算 $\boldsymbol{q}$.
   （2）对某较大的数 k ，计算 P^k. 再利用 P^k 的一列当作 $\boldsymbol{q}$ 的近似.

   （学习指导中描述了一个程序 nulbasis，它几乎使方法（1）自动化.）

   利用 $k=100$ 或更大的数，在你的矩阵程序允许的范围内，对一个随意给定的最大的随机矩阵做实验. 对每一种方法，给出你开始按键和运行程序所需要的时间. （MATLAB 的某些版本可用命令 flops 和 tic…toc 来记录浮点运算的次数和使用 MATLAB 计算总的消耗时间.）对比每个方法的优势并说出你倾向于使用哪一种方法.

练习题答案

1. a. 因为在一年之内有 5%的城市居民将迁移到郊区，每个人有 5%的这样的选择的可能性. 关于这个人没有更进一步的说明，所以我们说这个人去郊区的可能性是 5%. 这个事实蕴涵在状态向量 $\boldsymbol{x}_1$ 的第二个数值中，这里

$$\boldsymbol{x}_1 = M\boldsymbol{x}_0 = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.95 \\ 0.05 \end{bmatrix}$$

   b. 两年之后这个人住在郊区的可能性是 9.6%，这是因为

$$\boldsymbol{x}_2 = M\boldsymbol{x}_1 = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix}\begin{bmatrix} 0.95 \\ 0.05 \end{bmatrix} = \begin{bmatrix} 0.904 \\ 0.096 \end{bmatrix}$$

2. 稳态向量满足 $P\boldsymbol{x}=\boldsymbol{x}$ ，因为

$$P\boldsymbol{q} = \begin{bmatrix} 0.6 & 0.2 \\ 0.4 & 0.8 \end{bmatrix}\begin{bmatrix} 0.3 \\ 0.7 \end{bmatrix} = \begin{bmatrix} 0.32 \\ 0.68 \end{bmatrix} \neq \boldsymbol{q}$$

   我们断定 $\boldsymbol{q}$不是P 的稳态向量.

3. 例 1 中的 M 是一个正则的随机矩阵，这是因为它的数值全是严格正的. 所以可以使用定理 18，由例 4 我们已经知道其稳态向量，所以人口分布向量 $\boldsymbol{x}_k$ 趋向

$$\boldsymbol{q} = \begin{bmatrix} 0.375 \\ 0.625 \end{bmatrix}$$

   最终有 62.5%的人口将住在郊区.

补充习题

1. 对每个命题，标出是对还是错，验证每个答案.（若是对的，给出恰当的事实或定理；若是错的，解释原因并给出反例，说明为什么该命题在每种情况下都是错的.）在（a）～（f）中，$\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 是非零有限维向量空间 V 中的向量，而且 $S=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$.
   a. $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 的所有线性组合所成的集合是一个向量空间.
   b. 若 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_{p-1}\}$ 生成 V ，则 S 生成 V .
   c. 若 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_{p-1}\}$ 是线性无关的，则 S 也是线性无关的.
   d. 若 S 是线性无关的，则 S 是 V 的一组基.
   e. 若 $\mathrm{Span}\,S=V$ ，则 S 的某一子集是 V 的一组基.
   f. 若 $\dim V=p$ 且 $\mathrm{Span}\,S=V$ ，则 S 不能是线性

相关的.

g. $\mathbb{R}^3$ 中的一个平面是一个二维子空间.

h. 矩阵中的非主元列总是线性相关的.

i. 矩阵 A 中的行变换能够改变 A 的行之间的线性相关关系.

j. 矩阵中的行变换能够改变零空间.

k. 矩阵的秩等于非零行的个数.

l. 若 $m\times n$ 矩阵 A 行等价于阶梯形矩阵 U ，U 有 k 个非零行，则 $A\boldsymbol{x}=\boldsymbol{0}$ 的解空间的维数是 $m-k$.

m. 若 B 是由 A 经过几次行初等变换得到，则 $\text{rank}\, B=\text{rank}\, A$.

n. 矩阵 A 的非零行构成 $\text{Row}\, A$ 的一组基.

o. 若矩阵 A 和 B 有相同的简化阶梯形，则 $\text{Row}\, A=\text{Row}\, B$.

p. 若 H 是 $\mathbb{R}^3$ 的子空间，则存在一个 3×3 矩阵 A 使得 $H=\text{Col}\, A$.

q. 若 A 是 $m\times n$ 矩阵且 $\text{rank}\, A=m$ ，则线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 是唯一的.

r. 若 A 是 $m\times n$ 矩阵且线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 是射上的，则 $\text{rank}\, A=m$.

s. 坐标变换矩阵总是可逆的.

t. 若 $\mathcal{B}=\{\boldsymbol{b}_1,\ldots,\boldsymbol{b}_n\}$ 和 $\mathcal{C}=\{\boldsymbol{c}_1,\ldots,\boldsymbol{c}_n\}$ 是向量空间 V 的基，则坐标变换矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的第 j 列是坐标向量 $[\boldsymbol{c}_j]_{\mathcal{B}}$.

2. 求所有形如下列形式的向量构成集合的一组基.（计算要仔细.）

$$\begin{bmatrix} a-2b+5c \\ 2a+5b-8c \\ -a-4b+7c \\ 3a+b+c \end{bmatrix}$$

3. 令 $\boldsymbol{u}_1=\begin{bmatrix}-2\\4\\-6\end{bmatrix}$, $\boldsymbol{u}_2=\begin{bmatrix}1\\2\\-5\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}$, $W=\text{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$，求 W 的一个隐式的表达；即求一个或多个能刻画 W 中点的齐次方程的集合.（提示：$\boldsymbol{b}$ 何时在 W 中？）

4. 解释说明下面的讨论哪里有错：设 $\boldsymbol{f}(t)=3+t$, $\boldsymbol{g}(t)=3t+t^2$ ，因为 $\boldsymbol{g}(t)=t\boldsymbol{f}(t)$ ，$\boldsymbol{g}$ 是 $\boldsymbol{f}$ 的倍数，所以 $\{\boldsymbol{f},\boldsymbol{g}\}$ 是线性相关的.

5. 考虑多项式 $\boldsymbol{p}_1(t)=1+t$ ，$\boldsymbol{p}_2(t)=1-t$ ，$\boldsymbol{p}_3(t)=4$ ，$\boldsymbol{p}_4(t)=t+t^2$ ，$\boldsymbol{p}_5(t)=1+2t+t^2$ ，并设 H 是由集合 $S=\{\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3,\boldsymbol{p}_4,\boldsymbol{p}_5\}$ 所生成的 $\mathbb{P}_5$ 的子空间，使用（4.3 节）生成集定理证明中给出的方法求 H 的一组基.（说明如何从 S 中选取适当的向量.）

6. 假设 $\boldsymbol{p}_1$,$\boldsymbol{p}_2$,$\boldsymbol{p}_3$,$\boldsymbol{p}_4$ 是给定的多项式，它们生成 $\mathbb{P}_5$ 的一个二维子空间 H . 试叙述通过检查这 4 个多项式，你如何能几乎不用计算就能求出 H 的一组基.

7. 关于一个具有 18 个线性方程 20 个变量的齐次方程组的解集，为了知道每一个相应的非齐次方程组有解，你必须要知道什么？讨论之.

8. 令 H 是 n 维向量空间 V 的一个 n 维子空间，解释为什么 $H=V$.

9. 设 $T:\mathbb{R}^n\to\mathbb{R}^m$ 是线性变换，

a. 如果 T 是一对一映射，那么 T 的值域的维数是多少？解释之.

b. 如果 T 是从 $\mathbb{R}^n$ 映上到 $\mathbb{R}^m$，那么 T 的核（见 4.2 节）的维数是多少？解释之.

10. 令 S 是向量空间 V 中的一个最大线性无关子集，即 S 有这样的性质：若一个向量不在 S 中，添加到 S 后得到一个新的集合，则新的集合不再是线性无关的. 证明：S 一定是 V 的一组基.（提示：若 S 是线性无关的，但不是 V 的一组基，则 S 将如何？）

11. 令 S 是向量空间 V 的一个有限极小生成集，即，S 有这样的性质：若从 S 中去掉一个向量，则新的集合将不再生成 V ，证明：S 一定是 V 的一组基.

12~17 题给出了有时在应用中用到的秩的性质. 假设矩阵 A 是 $m\times n$ 矩阵.

12. 由（a）和（b），证明：$\text{rank}\, AB$ 不能超过 A 的秩或 B 的秩.（一般而言，矩阵乘积的秩不能

超过任何一个因子的秩.)

a. 若 B 是 $n\times p$ 矩阵，证明：$\text{rank}\,AB\leqslant\text{rank}\,A$.

（提示：解释为什么 AB 的列空间中的每个向量属于 A 的列空间中.）

b. 若 B 是 $n\times p$ 矩阵，证明：$\text{rank}\,AB\leqslant\text{rank}\,B$.

（提示：利用（a）研究 $\text{rank}(AB)^{\text{T}}$.）

13. 若 P 是 $m\times m$ 可逆矩阵，证明：$\text{rank}\,PA=\text{rank}\,A$.（提示：对 PA 和 $P^{-1}(PA)$ 应用第 12 题.）

14. 若 Q 是可逆的，证明：$\text{rank}\,AQ=\text{rank}\,A$.（提示：用第 13 题来研究 $\text{rank}(AQ)^{\text{T}}$.）

15. 设 A 是 $m\times n$ 矩阵，B 是 $n\times p$ 矩阵，且 $AB=0$. 证明：$\text{rank}\,A+\text{rank}\,B\leqslant n$.（提示：$\text{Nul}\,A$, $\text{Col}\,A$, $\text{Nul}\,B$, $\text{Col}\,B$ 四个子空间中有一个被包含在另外三个中的其中一个子空间.）

16. 设 A 是秩为 r 的 $m\times n$ 矩阵，则 A 的秩分解是具有 $A=CR$ 形式的等式，这里 C 是秩为 r 的 $m\times r$ 矩阵，R 是秩为 r 的 $r\times n$ 矩阵. 这样的分解总是存在的（参见 4.6 节习题 38）. 任给两个 $m\times n$ 矩阵 A 和 B，利用秩分解证明：

$$\text{rank}(A+B)\leqslant\text{rank}\,A+\text{rank}\,B$$

（提示：把 $A+B$ 写成两个分块矩阵的积.）

17. 矩阵 A 的**子矩阵**是由去掉 A 中若干行若干列（或不去掉）得到的矩阵. 可以证明 A 的秩为 r 当且仅当 A 包含一个 $r\times r$ 可逆子矩阵同时没有更大的可逆子方阵. 通过解释下面的（a）和（b）证明这个命题的一部分:（a）为什么秩为 r 的 $m\times n$ 矩阵 A 有一个秩为 r 的 $m\times r$ 子矩阵 A_1，（b）为什么 A_1 有一个可逆的 $r\times r$ 子矩阵 A_2.

秩这个概念在工程控制系统的设计中起着重要的作用，比如在本章引例中提到的航天飞机系统. 控制系统的状态空间模型包括一个形如

$$\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k+B\boldsymbol{u}_k\quad k=0,1,\cdots \qquad (1)$$

的差分方程，这里 A 是 $n\times n$ 矩阵，B 是 $n\times m$ 矩阵，$\{\boldsymbol{x}_k\}$ 是 $\mathbb{R}^n$ 中的一个“状态向量”序列，它用来描述在离散时间上系统的状态，$\{\boldsymbol{u}_k\}$ 是一个控制或输入序列. (A,B) 称为可控制的，如果

$$\text{rank}[B\ \ AB\ \ A^2B\ \cdots A^{n-1}B]=n \qquad (2)$$

（2）式中的矩阵称为系统的可控制性矩阵. 若 (A,B) 是可控制的，则通过简单地选择一个适当的 $\mathbb{R}^m$ 中的控制序列，最多在 n 步内，该系统可以被控制或被驱动，从状态 $\mathbf{0}$ 去到任一指定的（$\mathbb{R}^n$ 中的）状态 $\boldsymbol{v}$. 对这一点，习题 18 给出了在 $n=4$，$m=2$ 时的示例. 对可控制性的进一步讨论，可参见本教材的网站（第 4 章的案例研究）.

18. 设 A 是 4×4 矩阵，B 是 4×2 矩阵，$\boldsymbol{u}_0,\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3,\boldsymbol{u}_4$ 是 $\mathbb{R}^2$ 中的输入向量序列.

a. 设 $\boldsymbol{x}_0=\mathbf{0}$，用（1）式计算 $\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3,\boldsymbol{x}_4$，用（2）式中的可控制矩阵 M 写一个计算 $\boldsymbol{x}_4$ 的公式.（注意：矩阵 M 用分块矩阵来构造. 此处它的大小是 4×8.）

b. 设 (A,B) 是可控制的，$\boldsymbol{v}$ 是 $\mathbb{R}^4$ 中的任一向量，说明为何存在 $\mathbb{R}^2$ 中的一个控制序列 $\boldsymbol{u}_0,\cdots,\boldsymbol{u}_4$，使得 $\boldsymbol{x}_4=\boldsymbol{v}$.

在 19~22 题中，判定矩阵对是否是可控制的.

19. $A=\begin{bmatrix}0.9&1&0\\0&-0.9&0\\0&0&0.5\end{bmatrix}$，$B=\begin{bmatrix}0\\1\\1\end{bmatrix}$

20. $A=\begin{bmatrix}0.8&-0.3&0\\0.2&0.5&1\\0&0&-0.5\end{bmatrix}$，$B=\begin{bmatrix}1\\1\\0\end{bmatrix}$

21. [M] $A=\begin{bmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\-2&-4.2&-4.8&-3.6\end{bmatrix}$，$B=\begin{bmatrix}1\\0\\0\\-1\end{bmatrix}$

22. [M] $A=\begin{bmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\-1&-13&-12.2&-1.5\end{bmatrix}$，$B=\begin{bmatrix}1\\0\\0\\-1\end{bmatrix}$

第5章　特征值与特征向量

介绍性实例　动力系统与斑点猫头鹰

1990 年，在利用或滥用太平洋西北部大面积森林问题上，北方的斑点猫头鹰成为一个争论的焦点．环境保护学家试图说服联邦政府，如果采伐原始森林（长有 200 年以上的树木）的行为得不到制止的话，猫头鹰将濒临灭绝的危险，因为猫头鹰喜好在那里居住．而木材行业却争辩说猫头鹰不应被划为“濒临灭绝动物”，并引用一些已发表的科学报告来支持其论点[㊀]．对木材行业来说，如果政府出台新的伐木限制的话，预计将失去 30 000 至 100 000 个工作岗位．

数学生态学家们处于争论双方的中间，由于争论的双方都想游说数学家们，数学生态学家们加快了他们对斑点猫头鹰种群的动力学研究．猫头鹰的生命周期自然分为三个阶段：幼年期（1 岁以前）、半成年期（1~2 岁）、成年期（2 岁以后）．猫头鹰交配在半成年和成年期，开始生育繁殖，可活到 20 岁左右．每一对猫头鹰需要约 1000 公顷(4 平方英里）的土地作为自己的栖息地．生命周期的关键期是当幼年猫头鹰离开巢的时候，为生存和进入半成年期，一只幼年猫头鹰必须成功地找到一个新的栖息地安家（通常还带有一个配偶）．

研究种群动力学的第 1 步是建立以每年的种群量为区间的种群模型，时间为 $k=0,1,2,\cdots$.通常可以假设在每一个生命阶段雄性和雌性的比例为 1:1，而且只计算雌性猫头鹰，第 k 年的种群量可以用向量 $\boldsymbol{x}_k=(j_k,s_k,a_k)$ 表示，其中 j_k,s_k 和 a_k 分别代表雌性猫头鹰在幼年期、半成年期和成年期的数量．

利用人口统计研究的实际现场数据，R. Lamberson 及其同事设计了下面的“阶段矩阵模型”(stage - matrix　model)[㊁]：

$$\begin{bmatrix} j_{k+1} \\ s_{k+1} \\ a_{k+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0.33 \\ 0.18 & 0 & 0 \\ 0 & 0.71 & 0.94 \end{bmatrix} \begin{bmatrix} j_k \\ s_k \\ a_k \end{bmatrix}$$

㊀ “The Great Spotted Owl War”，《读者文摘》，November 1992，pp. 91-95.

㊁ R.H.Lamberson, R.Mckelvey, B.R.Noon, and C.Voss, “A Dynamic Analysis of the Viability of the Northern Spotted Owl in a Fragmented Forest Environment”，*Conservation Biology* 6（1992)，505-512，还可参见 Lamberson 教授的私人通信，1993.

在这里新的幼年雌性猫头鹰在 $k+1$ 年中的数量是成年雌性猫头鹰在 k 年里数量的 0.33 倍（根据每一对猫头鹰的平均生殖率而定）. 此外，18%的幼年雌性猫头鹰得以生存进入半成年期，71%的半成年雌性猫头鹰和 94%的成年雌性猫头鹰生存下来被计为成年猫头鹰.

阶段矩阵模型是形式为 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 的差分方程，这种方程被称为**动力系统**（或**离散线性动力系统**），因为它描述的是系统随时间推移的变化.

Lamberson 阶段矩阵中，18%的幼年猫头鹰生存率是受原始森林中猫头鹰数目影响最大的项目. 事实上，60%的幼年猫头鹰通常生存下来后就会离开自己的巢，但是在 Lamberson 和他的同事们作研究的加州的 Willow Creek 地区，只有 30%的幼年猫鹰在弃巢后能找到新的栖息地，其他的在寻找新家园过程中失踪了.

猫头鹰不能找到新栖息地的一个重要原因是因为对原始森林分散区域的砍伐而加剧了原始森林的分割. 当猫头鹰离开森林保护区并穿过一块滥伐地时，被捕食动物袭击的危险大增. 5.6 节将会展示前面讨论的模型如何预测斑点猫头鹰的最终灭绝，但如果 50%的幼年猫头鹰弃巢后能找到新的栖息地，猫头鹰种群将会兴旺起来.

本章的目的是剖析线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的作用，把它分解为容易理解的元素. 除了 5.4 节，整章中出现的矩阵都是方阵. 虽然这里讨论的主要应用是离散动力系统（包括上面的斑点猫头鹰）. 但有关特征值和特征向量的基本概念对纯数学和应用数学都很有用，它们出现的背景要比我们这里考虑的要广泛得多. 同样，特征值还被用来研究微分方程和连续的动力系统，为工程设计提供关键知识，自然地，也出现在物理和化学等领域里.

5.1 特征向量与特征值

尽管变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 有可能使向量往各个方向移动，但通常会有某些特殊向量，A 对这些向量的作用是很简单的.

例 1 设 $A=\begin{bmatrix}3 & -2\\ 1 & 0\end{bmatrix}$, $\boldsymbol{u}=\begin{bmatrix}-1\\ 1\end{bmatrix}$, $\boldsymbol{v}=\begin{bmatrix}2\\ 1\end{bmatrix}$. 图 5-1 显示 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 乘以 A 后的图像. 事实上，$A\boldsymbol{v}$ 正好是 $2\boldsymbol{v}$，因此，A 仅仅是“拉伸”了 $\boldsymbol{v}$. ■

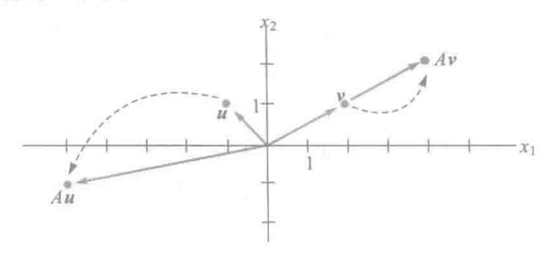


图 5-1 乘以 A 的作用

作为另一个例子，4.9 节的读者应该能回忆起，假如 A 是随机矩阵，那么 A 的稳态向量 $\boldsymbol{q}$ 满足方程 $A\boldsymbol{x}=\boldsymbol{x}$，即 $A\boldsymbol{q}=1\cdot\boldsymbol{q}$.

这一节，我们将研究形如 $Ax = 2x$ 或 $Ax = -4x$ 的方程，并且去寻找那些被 A 变换成自身一个数量倍的向量.

定义 A 为 $n \times n$ 矩阵，x 为非零向量，若存在数 λ 使 $Ax = \lambda x$ 成立，则称 λ 为 A 的特征值，x 称为对应于 λ 的特征向量.[⊖]

容易验证给定的向量是否是矩阵的特征向量，也容易判断给出的数是否是特征值.

例 2 设 $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$，$u = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$ 和 $v = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$，$u$ 和 v 是否是 A 的特征向量？

解

$$Au = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} 6 \\ -5 \end{bmatrix} = \begin{bmatrix} -24 \\ 20 \end{bmatrix} = -4\begin{bmatrix} 6 \\ -5 \end{bmatrix} = -4u$$

$$Av = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 11 \end{bmatrix} \neq \lambda\begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

因此，u 是特征值 -4 对应的特征向量，但 Av 不是 v 的倍数（见图 5-2)，故 v 不是 A 的特征向量.

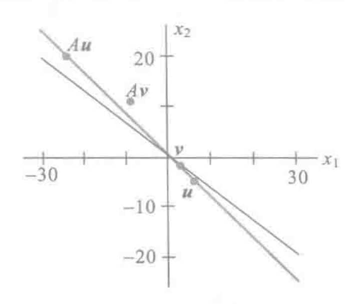


图 5-2 $Au=-4u$，但 $Av \neq \lambda v$ ■

例 3 证明 7 是例 2 中矩阵 A 的特征值，并求特征值 7 对应的特征向量.

解 数 7 是 A 的特征值当且仅当方程

$$Ax = 7x \tag{1}$$

有非平凡解，（1）等价于 $Ax - 7x = 0$，或

$$(A - 7I)x = 0 \tag{2}$$

为解该齐次方程，计算

$$A - 7I = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix}$$

$A - 7I$ 的列显然是线性相关的，故（2）有非平凡解，因此 7 是 A 的特征值. 为求其对应的特征向量，用行变换化简矩阵：

$$\begin{bmatrix} -6 & 6 & 0 \\ 5 & -5 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

故通解为 $x_2\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. 凡是具有此种形式且 $x_2 \neq 0$ 的向量都是 $\lambda = 7$ 对应的特征向量. ■

⊖ 注意，根据定义，特征向量必须是非零的，但特征值可以为零. 特征值为 0 的情况将在例 5 后面讨论.

警告 虽然在例 3 中使用了行化简来求特征向量，但不能用行化简求特征值.矩阵 A 的阶梯形通常不显示出 A 的特征值.

方程（1）和（2）的等价性对任何 $\lambda=7$ 以外的λ显然都是成立的. 故 λ 是 A 的特征值当且仅当方程

$$(A-\lambda I)\boldsymbol{x}=\mathbf{0} \tag{3}$$

有非平凡解. 方程(3)的所有解的集合就是矩阵 $A-\lambda I$ 的零空间. 因此，该集合是 $\mathbb{R}^n$ 的子空间，称为 A 的对应于 λ 的特征空间. 特征空间由零向量和所有对应于 λ 的特征向量组成.

从例 3 看出，例 2 矩阵 A 对应 $\lambda=7$ 的特征空间由（1, 1）的所有倍数 x_2（1, 1）组成，因此，特征空间是过（1, 1）和原点的直线. 在例 2，也可以验证对应 $\lambda=-4$ 的特征空间是经过（6, −5）和原点的直线. 图 5-3 显示了这些特征空间、特征向量（1, 1）和（3/2, −5/4）及变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 对每个特征空间的几何意义.

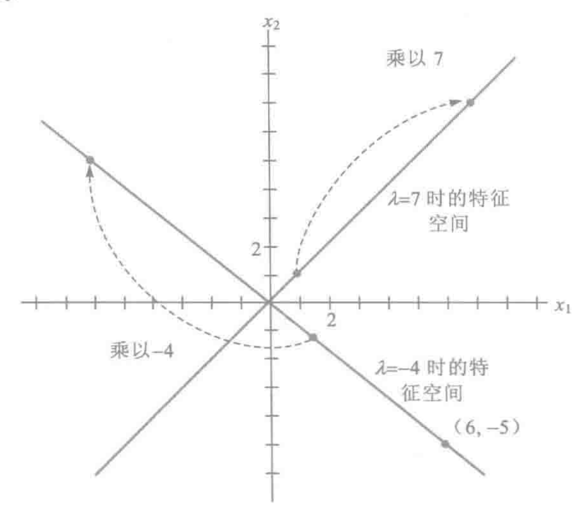


图 5-3 对应 $\lambda=-4$ 和 $\lambda=7$ 的特征空间

例 4 设 $A=\begin{bmatrix}4 & -1 & 6\\ 2 & 1 & 6\\ 2 & -1 & 8\end{bmatrix}$，$A$ 的一个特征值是 2，求对应的特征空间的一个基.

解 计算

$$A-2I=\begin{bmatrix}4 & -1 & 6\\ 2 & 1 & 6\\ 2 & -1 & 8\end{bmatrix}-\begin{bmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{bmatrix}=\begin{bmatrix}2 & -1 & 6\\ 2 & -1 & 6\\ 2 & -1 & 6\end{bmatrix}$$

用行变换化简 $(A-2I)\boldsymbol{x}=\mathbf{0}$ 的增广矩阵：

$$\begin{bmatrix}2 & -1 & 6 & 0\\ 2 & -1 & 6 & 0\\ 2 & -1 & 6 & 0\end{bmatrix}\sim\begin{bmatrix}2 & -1 & 6 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}$$

因为方程 $(A-2I)\boldsymbol{x}=\mathbf{0}$ 有自由变量，故 2 是 A 的特征值. 通解是

$$\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=x_2\begin{bmatrix}1/2\\1\\0\end{bmatrix}+x_3\begin{bmatrix}-3\\0\\1\end{bmatrix},\ x_2\text{和}x_3\text{为任意值}$$

图 5-4 显示出特征空间是 $\mathbb{R}^3$ 的 2 维子空间，其中的一个基是

$$\left\{\begin{bmatrix}1\\2\\0\end{bmatrix},\begin{bmatrix}-3\\0\\1\end{bmatrix}\right\}$$

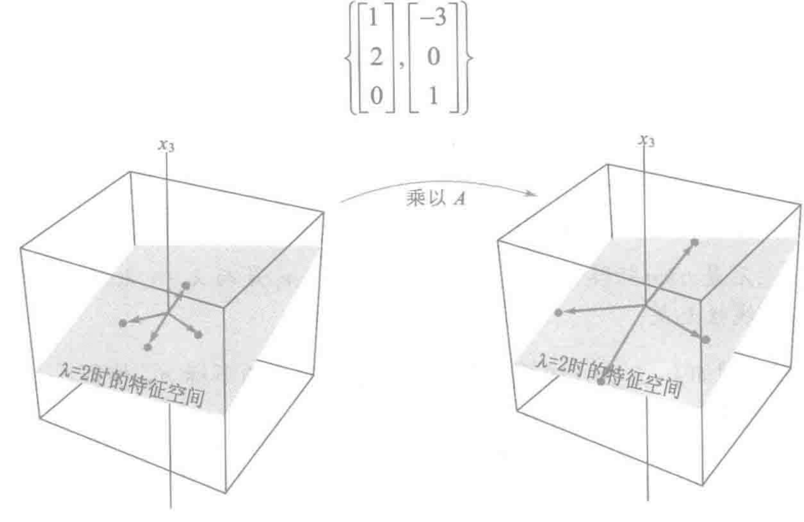


图 5-4　A 对特征空间的扩张作用　■

数值计算的注解　在已知特征值的条件下，例 4 提供了手工计算特征向量的方法. 虽然也可以利用矩阵程序和行化简来找出特征空间，但这不完全可靠. 舍入误差有时会使简化的阶梯形矩阵出现错误的主元素. 如果矩阵不是很大，最好的计算机程序可按要求的精度同时算出特征值和特征向量的近似值. 随着计算能力和软件的改善，能被分析的矩阵规模也逐年增大.

下列定理描述了特征值能被准确求出的几种特例之一. 有关特征值的计算将在下一节讨论.

定理 1　三角矩阵的主对角线的元素是其特征值.

证　为简单起见，考虑 3×3 的情形. 假设 A 是上三角矩阵，那么 $A-\lambda I$ 是：

$$A-\lambda I=\begin{bmatrix}a_{11}&a_{12}&a_{13}\\0&a_{22}&a_{23}\\0&0&a_{33}\end{bmatrix}-\begin{bmatrix}\lambda&0&0\\0&\lambda&0\\0&0&\lambda\end{bmatrix}$$
$$=\begin{bmatrix}a_{11}-\lambda&a_{12}&a_{13}\\0&a_{22}-\lambda&a_{23}\\0&0&a_{33}-\lambda\end{bmatrix}$$

数 λ 是 A 的特征值当且仅当方程 $(A-\lambda I)\boldsymbol{x}=\mathbf{0}$ 有非平凡解，即方程有自由变量. 由 $A-\lambda I$ 容易看出，$(A-\lambda I)\boldsymbol{x}=\mathbf{0}$ 有自由变量当且仅当 $A-\lambda I$ 的对角线上的元素至少有一个为零，而当 λ 取 a_{11}, a_{22},

a_{33} 之一的时候出现这种情况.故 A 的特征值是 a_{11}, a_{22}, a_{33}. 若 A 是下三角矩阵，见习题 28. ■

例 5 设 $A=\begin{bmatrix}3 & 6 & -8\\0 & 0 & 6\\0 & 0 & 2\end{bmatrix}, B=\begin{bmatrix}4 & 0 & 0\\-2 & 1 & 0\\5 & 3 & 4\end{bmatrix}$，$A$ 的特征值是 3，0 和 2. B 的特征值是 4 和 1. ■

例 5 中的矩阵 A 有零特征值，因此方程

$$Ax = 0x \tag{4}$$

有非平凡解. 但（4）等价于 $Ax=\mathbf{0}$，而 $Ax=\mathbf{0}$ 有非平凡解的充要条件是 A 是不可逆的. 因此，A 有零特征值的充要条件是 A 不可逆. 由此，0 是 A 的特征值当且仅当 A 不可逆. 这一事实添加到 5.2 节的可逆矩阵定理中.

接下来的定理很重要，以后将会用到. 定理的证明展示了特征向量的典型计算.

定理 2 $\lambda_1,\cdots,\lambda_r$ 是 $n\times n$ 矩阵 A 相异的特征值，$\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r$ 是与 $\lambda_1,\cdots,\lambda_r$ 对应的特征向量，那么向量集合 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r\}$ 线性无关.

证 用反证法. 假如 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r\}$ 线性相关，那么存在最小的下标 p，使得 $\boldsymbol{v}_{p+1}$ 是前面向量（这些向量线性无关）的线性组合，即存在数 $c_1,\cdots,c_p$，使得

$$c_1\boldsymbol{v}_1+\cdots+c_p\boldsymbol{v}_p=\boldsymbol{v}_{p+1} \tag{5}$$

式（5）两边乘 A，并将 $A\boldsymbol{v}_k=\lambda_k\boldsymbol{v}_k (k=1,\cdots,p+1)$ 代入，得

$$c_1A\boldsymbol{v}_1+\cdots+c_pA\boldsymbol{v}_p=A\boldsymbol{v}_{p+1}$$

$$c_1\lambda_1\boldsymbol{v}_1+\cdots+c_p\lambda_p\boldsymbol{v}_p=\lambda_{p+1}\boldsymbol{v}_{p+1} \tag{6}$$

式（5）两边乘 λ_{p+1}，并与式（6）相减，得

$$c_1(\lambda_1-\lambda_{p+1})\boldsymbol{v}_1+\cdots+c_p(\lambda_p-\lambda_{p+1})\boldsymbol{v}_p=\mathbf{0} \tag{7}$$

因为 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 线性无关，式（7）的系数全为零. 但 $\lambda_i-\lambda_{p+1},\cdots,\lambda_p-\lambda_{p+1}$ 全不为零，因此只有 $c_i=0, i=1,\cdots,p$. 由式(5)得 $\boldsymbol{v}_{p+1}=\mathbf{0}$，矛盾. 故 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r\}$ 必是线性无关的. ■

特征向量与差分方程

我们通过构造一阶差分方程

$$\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k \quad (k=0,1,2,\cdots) \tag{8}$$

的解来结束本节.

若 A 是 $n\times n$ 矩阵，那么方程（8）是 $\mathbb{R}^n$ 的序列 $\{\boldsymbol{x}_k\}$ 的递归表示. 方程（8）的解是描述序列 $\{\boldsymbol{x}_k\}$ 的每个 $\boldsymbol{x}_k$ 的显式公式，公式不直接依赖于 A 和序列前面的项，而是依赖于初始项 $\boldsymbol{x}_0$.

构造方程（8）的解的最简单方法是取 A 的一个特征向量 $\boldsymbol{x}_0$ 和它对应的特征值 λ，然后令

$$\boldsymbol{x}_k=\lambda^k\boldsymbol{x}_0 \quad (k=1,2,\cdots) \tag{9}$$

这就是方程（8）的解，因为

$$A\boldsymbol{x}_k=A(\lambda^k\boldsymbol{x}_0)=\lambda^k(A\boldsymbol{x}_0)=\lambda^k(\lambda\boldsymbol{x}_0)=\lambda^{k+1}\boldsymbol{x}_0=\boldsymbol{x}_{k+1}$$

此外，形如（9）的解的线性组合仍然是（9）的解，见习题 33.

练习题

1. $A=\begin{bmatrix}6 & -3 & 1\\ 3 & 0 & 5\\ 2 & 2 & 6\end{bmatrix}$,5 是 A 的特征值吗？

2. 若 $\boldsymbol{x}$ 是 A 对应于λ的特征向量，求 $A^3\boldsymbol{x}$.

3. 假设 $\boldsymbol{b}_1$ 和 $\boldsymbol{b}_2$ 分别是对应于不同的特征值 λ_1 和 λ_2 的特征向量，$\boldsymbol{b}_3$ 和 $\boldsymbol{b}_4$ 是对应于第三个不同的特征值 λ_3 的线性无关特征向量. $\{\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3, \boldsymbol{b}_4\}$ 一定是线性无关集吗？（**提示**：考虑方程 $c_1\boldsymbol{b}_1+c_2\boldsymbol{b}_2+(c_3\boldsymbol{b}_3+c_4\boldsymbol{b}_4)=\boldsymbol{0}$.）

习题 5.1

1. $\lambda=2$ 是 $\begin{bmatrix}3 & 2\\ 3 & 8\end{bmatrix}$ 的特征值吗？为什么？

2. $\lambda=-3$ 是 $\begin{bmatrix}-1 & 4\\ 6 & 9\end{bmatrix}$ 的特征值吗？为什么？

3. $\begin{bmatrix}1\\ 3\end{bmatrix}$ 是 $\begin{bmatrix}1 & -1\\ 6 & -4\end{bmatrix}$ 的特征向量吗？如果是，求对应的特征值.

4. $\begin{bmatrix}-1\\ 1\end{bmatrix}$ 是 $\begin{bmatrix}5 & 2\\ 3 & 6\end{bmatrix}$ 的特征向量吗？如果是，求对应的特征值.

5. $\begin{bmatrix}3\\ -2\\ 1\end{bmatrix}$ 是 $\begin{bmatrix}-4 & 3 & 3\\ 2 & 3 & -2\\ -1 & 0 & -2\end{bmatrix}$ 的特征向量吗？如果是，求对应的特征值.

6. $\begin{bmatrix}1\\ -2\\ 2\end{bmatrix}$ 是 $\begin{bmatrix}3 & 6 & 7\\ 3 & 2 & 7\\ 5 & 6 & 4\end{bmatrix}$ 的特征向量吗？如果是，求对应的特征值.

7. $\lambda=4$ 是 $\begin{bmatrix}3 & 0 & -1\\ 2 & 3 & 1\\ -3 & 4 & 5\end{bmatrix}$ 的特征值吗？如果是，求λ对应的一个特征向量.

8. $\lambda=1$ 是 $\begin{bmatrix}4 & -2 & 3\\ 0 & -1 & 3\\ -1 & 2 & -2\end{bmatrix}$ 的特征值吗？如果是，求 λ 对应的一个特征向量.

在习题 9~16 中，求所给特征值对应的特征空间的一个基.

9. $A=\begin{bmatrix}3 & 0\\ 2 & 1\end{bmatrix}, \lambda=1,3$

10. $A=\begin{bmatrix}4 & 2\\ 3 & 1\end{bmatrix}, \lambda=-5$

11. $A=\begin{bmatrix}1 & -3\\ -4 & 5\end{bmatrix}, \lambda=-1,7$

12. $A=\begin{bmatrix}4 & 1\\ 3 & 6\end{bmatrix}, \lambda=3,7$

13. $A=\begin{bmatrix}4 & 0 & 1\\ -2 & 1 & 0\\ -2 & 0 & 1\end{bmatrix}, \lambda=1,2,3$

14. $A=\begin{bmatrix}4 & 0 & -1\\ 3 & 0 & 3\\ 2 & -2 & 5\end{bmatrix}, \lambda=3$

15. $A=\begin{bmatrix}-4 & 1 & 1\\ 2 & -3 & 2\\ 3 & 3 & -2\end{bmatrix}, \lambda=-5$

16. $A=\begin{bmatrix}5 & 0 & -1 & 0\\ 1 & 3 & 0 & 0\\ 2 & -1 & 3 & 0\\ 4 & -2 & -2 & 4\end{bmatrix}, \lambda=4$

求习题 17 和习题 18 中矩阵的特征值.

17. $\begin{bmatrix}0 & 0 & 0\\ 0 & 3 & 4\\ 0 & 0 & -2\end{bmatrix}$

18. $\begin{bmatrix}5 & 0 & 0\\ 0 & 0 & 0\\ -1 & 0 & 3\end{bmatrix}$

19. 不用计算，求 $A=\begin{bmatrix}1&2&3\\1&2&3\\1&2&3\end{bmatrix}$ 的一个特征值，验证你的结果.

20. 不用计算，求 $A=\begin{bmatrix}2&2&2\\2&2&2\\2&2&2\end{bmatrix}$ 的一个特征值和两个线性无关的特征向量，验证你的结果.

在习题 21 和习题 22 中，A 是 $n\times n$ 矩阵. 标出每个命题的真假，给出理由.

21. a. 若对某个向量 $\boldsymbol{x}$ 有 $A\boldsymbol{x}=\lambda\boldsymbol{x}$，则 λ 是 A 的特征值.
    b. 矩阵 A 不可逆的充要条件是零是 A 的特征值.
    c. 数 c 是 A 的特征值的充要条件是 $(A-cI)\boldsymbol{x}=\boldsymbol{0}$ 有非平凡解.
    d. 求 A 的特征向量可能是困难的，但验证一个给定的向量是否是特征向量却是容易的.
    e. 为求 A 的特征值，可将 A 化简为阶梯形矩阵.

22. a. 若存在数 λ 使 $A\boldsymbol{x}=\lambda\boldsymbol{x}$ 成立，那么 $\boldsymbol{x}$ 是 A 的特征向量.
    b. 若 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 是线性无关的特征向量，则它们是对应于不同特征值的特征向量.
    c. 随机矩阵的稳态向量就是其特征向量.
    d. 矩阵的特征值是在其主对角线上的元素.
    e. 矩阵 A 的特征空间是某矩阵的零空间.

23. 为什么 2×2 矩阵最多只能有 2 个相异的特征值？为什么 $n\times n$ 矩阵最多只能有 n 个相异的特征值？

24. 举一个只有一个特征值的 2×2 矩阵的例子.

25. 设 λ 是可逆矩阵 A 的特征值，证明 λ^{-1} 是 A^{-1} 的特征值. （提示：假设非零向量 $\boldsymbol{x}$ 满足 $A\boldsymbol{x}=\lambda\boldsymbol{x}$.）

26. 证明：若 A^2 是零矩阵，则 A 只有零特征值.

27. 证明：λ 是矩阵 A 的特征值当且仅当 λ 是 A^{T} 的特征值.（提示：找出 $A-\lambda I$ 和 $A^{\mathrm{T}}-\lambda I$ 的关系，然后证明 $A-\lambda I$ 可逆的充要条件是 $A^{\mathrm{T}}-\lambda I$ 是可逆的.）

28. 若 A 是下三角矩阵，利用习题 27 来证明定理 1.

29. 设 $n\times n$ 矩阵 A 的每行元素之和都等于 s，证明 s 是 A 的特征值.（提示：找一个特征向量.）

30. 设 $n\times n$ 矩阵 A 的每列元素之和都等于 s，证明 s 是 A 的特征值.（提示：利用习题 27 和 29.）

在习题 31 和习题 32 中，设 A 是线性变换 T 的矩阵，不写出 A，求 A 的特征值和特征空间.

31. 设 T 是 $\mathbb{R}^2$ 的相对过原点的某条直线的反射变换.

32. 设 T 是 $\mathbb{R}^3$ 的相对过原点的某条直线的旋转变换.

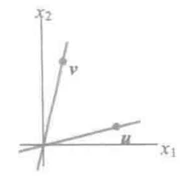
33. 设 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是矩阵 A 的特征值 λ 和 μ 对应的特征向量，c_1 和 c_2 是数，定义

$$\boldsymbol{x}_k=c_1\lambda^k\boldsymbol{u}+c_2\mu^k\boldsymbol{v}\quad(k=0,1,2,\cdots)$$

    a. 根据定义求 $\boldsymbol{x}_{k+1}$.
    b. 用 $\boldsymbol{x}_k$ 的公式计算 $A\boldsymbol{x}_k$，并验证 $A\boldsymbol{x}_k=\boldsymbol{x}_{k+1}$. 这就证明了上面定义的序列 $\{\boldsymbol{x}_k\}$ 满足差分方程 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k(k=0,1,2,\cdots)$.

34. 如果给出的初始向量 $\boldsymbol{x}_0$ 不是 A 的特征向量，你是否能找出解差分方程 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k(k=0,1,2,\cdots)$ 的方法？（提示：能否把 $\boldsymbol{x}_0$ 与 A 的特征向量相联系？）

35. 设 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是下图中的向量，并设 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是一个 2×2 矩阵 A 的分别对应于特征值 2 和 3 的特征向量. 并设 $T:\mathbb{R}^2\to\mathbb{R}^2$ 是线性变换，对 $\mathbb{R}^2$ 中的每一 $\boldsymbol{x}$ 有 $T(\boldsymbol{x})=A\boldsymbol{x}$，记 $\boldsymbol{w}=\boldsymbol{u}+\boldsymbol{v}$. 在下图中，仔细画出向量 $T(\boldsymbol{u})$，$T(\boldsymbol{v})$和 $T(\boldsymbol{w})$.



36. 设 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是一个 2×2 矩阵 A 分别对应特征值为 −1 和 3 的特征向量，重做 35 题.

[M]在习题 37~40 中，利用矩阵程序求矩阵的特征值. 然后用例 4 的行化简方法找出每一特征空

间的基.

37. $\begin{bmatrix} 12 & 1 & 4 \\ 2 & 11 & 4 \\ 1 & 3 & 7 \end{bmatrix}$

38. $\begin{bmatrix} 5 & -2 & 2 & -4 \\ 7 & -4 & 2 & -4 \\ 4 & -4 & 2 & 0 \\ 3 & -1 & 1 & -3 \end{bmatrix}$

39. $\begin{bmatrix} 12 & -90 & -30 & 30 & 30 \\ 8 & -49 & 15 & 15 & 15 \\ 16 & -52 & 12 & 0 & 20 \\ 0 & -30 & 10 & 22 & 10 \\ 8 & -41 & 15 & 15 & 7 \end{bmatrix}$

40. $\begin{bmatrix} -23 & 57 & -9 & -15 & -59 \\ -10 & 12 & -10 & 2 & -22 \\ 11 & 5 & -3 & -19 & -15 \\ -27 & 31 & -27 & 25 & -37 \\ -5 & -15 & -5 & 1 & 31 \end{bmatrix}$

练习题答案

1. 5 是 A 的特征值的充要条件是方程 $(A-5I)\boldsymbol{x}=\boldsymbol{0}$ 有非平凡解. 计算

$$A-5I=\begin{bmatrix} 6 & -3 & 1 \\ 3 & 0 & 5 \\ 2 & 2 & 6 \end{bmatrix}-\begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix}=\begin{bmatrix} 1 & -3 & 1 \\ 3 & -5 & 5 \\ 2 & 2 & 1 \end{bmatrix}$$

将增广矩阵行化简:

$$\begin{bmatrix} 1 & -3 & 1 & 0 \\ 3 & -5 & 5 & 0 \\ 2 & 2 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 1 & 0 \\ 0 & 4 & 2 & 0 \\ 0 & 8 & -1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 1 & 0 \\ 0 & 4 & 2 & 0 \\ 0 & 0 & -5 & 0 \end{bmatrix}$$

显然，齐次方程组无自由变量，故 $A-5I$ 是可逆矩阵，即 5 不是 A 的特征值.

2. 如果 $\boldsymbol{x}$ 是 A 对应 λ 的特征向量，则有 $A\boldsymbol{x}=\lambda\boldsymbol{x}$，故 $A^2\boldsymbol{x}=A(\lambda\boldsymbol{x})=\lambda A\boldsymbol{x}=\lambda^2\boldsymbol{x}$.
同理，$A^3\boldsymbol{x}=A(A^2\boldsymbol{x})=A(\lambda^2\boldsymbol{x})=\lambda^2 A\boldsymbol{x}=\lambda^3\boldsymbol{x}$.对一般的 k，有 $A^k\boldsymbol{x}=\lambda^k\boldsymbol{x}$.

3. 是. 假设 $c_1\boldsymbol{b}_1+c_2\boldsymbol{b}_2+c_3\boldsymbol{b}_3+c_4\boldsymbol{b}_4=\boldsymbol{0}$，因为相同特征值的特征向量的任意线性组合仍是该特征值的一个特征向量，$c_3\boldsymbol{b}_3+c_4\boldsymbol{b}_4$ 是 λ_3 的特征向量. 根据定理 2，向量 $\boldsymbol{b}_1$，$\boldsymbol{b}_2$ 和 $c_3\boldsymbol{b}_3+c_4\boldsymbol{b}_4$ 线性无关，所以 $c_1\boldsymbol{b}_1+c_2\boldsymbol{b}_2+(c_3\boldsymbol{b}_3+c_4\boldsymbol{b}_4)=\boldsymbol{0}$ 蕴涵 $c_1=c_2=0$. 但是由于 $\boldsymbol{b}_3$ 和 $\boldsymbol{b}_4$ 线性无关，c_3 和 c_4 必然为零. 因此，原始方程中的所有系数都为零，向量 $\boldsymbol{b}_1$，$\boldsymbol{b}_2$，$\boldsymbol{b}_3$ 和 $\boldsymbol{b}_4$ 线性无关.

5.2 特征方程

方阵 A 的特征方程是一个数量方程，其中包含了有关特征值的有用信息. 我们先从一个简单的例子开始，然后讨论一般情形.

例 1 求 $A=\begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$ 的特征值.

解 我们必须找出所有这样的数 λ，使得矩阵方程

$$(A-\lambda I)\boldsymbol{x}=\boldsymbol{0}$$

有非平凡解. 由逆矩阵定理，这个问题等价于要求出所有的 λ，使得矩阵 $A-\lambda I$ 为不可逆矩阵，这里

$$A-\lambda I=\begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}-\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}=\begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix}$$

由 2.2 节的定理 4，该矩阵当它的行列式值为零时是不可逆的. 因此 A 的特征值是下列方程的解

$$\det(A-\lambda I)=\det\begin{bmatrix}2-\lambda & 3\\ 3 & -6-\lambda\end{bmatrix}=0$$

我们知道

$$\det\begin{bmatrix}a & b\\ c & d\end{bmatrix}=ad-bc$$

故

$$\begin{aligned}\det(A-\lambda I)&=(2-\lambda)(-6-\lambda)-(3)(3)\\ &=-12+6\lambda-2\lambda+\lambda^2-9\\ &=\lambda^2+4\lambda-21\end{aligned}$$

令 $\lambda^2+4\lambda-21=0$，有 $(\lambda-3)(\lambda+7)=0$，故 A 的特征值是 3 和−7. ■

例 1 的行列式把包含 2 个未知数（λ 和 $\boldsymbol{x}$）的矩阵方程 $(A-\lambda I)\boldsymbol{x}=\boldsymbol{0}$ 转化为只含一个未知数的数值方程 $\lambda^2+4\lambda-21=0$，这种思想对 $n\times n$ 矩阵仍然适用. 但在讨论更大的矩阵之前，我们先给出研究特征值要用到的行列式的性质.

行列式

设 A 是 $n\times n$ 矩阵，U 是对 A 作行替换和行交换（不作行倍乘）所得到的任一阶梯形矩阵，r 是行交换的次数，那么 A 的**行列式** $\det A=(-1)^r u_{11}\cdots u_{nn}$. 如果 A 可逆，那么 $u_{11},\cdots,u_{nn}$ 都是主元. 否则，至少有 u_{nn} 为零，从而乘积 $u_{11}\cdots u_{nn}$ 为零. 因此[㊀]

$$\det A=\begin{cases}(-1)^r u_{11}\cdots u_{nn} & \text{当}A\text{可逆}\\ 0 & \text{当}A\text{不可逆}\end{cases}\tag{1}$$

例 2 计算 $A=\begin{bmatrix}1 & 5 & 0\\ 2 & 4 & -1\\ 0 & -2 & 0\end{bmatrix}$ 的行列式 $\det A$.

解 下面的行化简使用了一次行交换：

$$A\sim\begin{bmatrix}1 & 5 & 0\\ 0 & -6 & -1\\ 0 & -2 & 0\end{bmatrix}\sim\begin{bmatrix}1 & 5 & 0\\ 0 & -2 & 0\\ 0 & -6 & -1\end{bmatrix}\sim\begin{bmatrix}1 & 5 & 0\\ 0 & -2 & 0\\ 0 & 0 & -1\end{bmatrix}=U_1$$

故 $\det A=(-1)^1(1)(-2)(-1)=-2$. 下面选择的行化简没有用到行交换，得到了另一个不同的阶梯形矩阵.最后一步是用−1/3 乘第 2 行加到第 3 行上：

$$A\sim\begin{bmatrix}1 & 5 & 0\\ 0 & -6 & -1\\ 0 & -2 & 0\end{bmatrix}\sim\begin{bmatrix}1 & 5 & 0\\ 0 & -6 & -1\\ 0 & 0 & 1/3\end{bmatrix}=U_2$$

㊀ 公式（1）在 3.2 节导出，没有读过第 3 章的读者可能会把该公式当作行列式 $\det A$ 的定义. 值得注意的是，对 A 作变换（不作倍乘变换）所得到的任一阶梯形矩阵 U 的行列式值就是 $\det A$.

此时，$\det A=(-1)^0(1)(-6)(1/3)=-2$，与前面的计算结果相同. ■

计算行列式的公式（1）说明 A 是可逆的当且仅当 $\det A$ 非零. 这个事实以及 5.1 节中例 5 后面的可逆性的性质，可以添加到可逆矩阵定理当中.

定理 （可逆矩阵定理（续））

设 A 是 $n\times n$ 矩阵，则 A 是可逆的当且仅当

s. 0 不是 A 的特征值.

t. A 的行列式不等于零.

如果 A 是 3×3 矩阵，那么 $|\det A|$ 是由 A 的列向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 生成的平行六面体的体积，见图 5-5（细节见 3.3 节）. 平行六面体的体积不等于零当且仅当向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 线性无关，即矩阵 A 是可逆的.（如果向量非零且线性相关，它们位于同一个平面或一条直线上.）

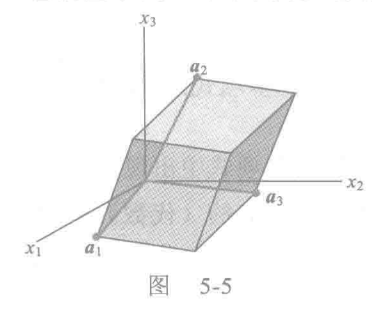


图 5-5

定理 3 列出了由 3.1 节和 3.2 节导出的结论，这里将（a）也包含进来以便参考.

定理 3 （行列式的性质）

设 A 和 B 是 $n\times n$ 矩阵.

a. A 可逆的充要条件是 $\det A\neq 0$.

b. $\det AB=(\det A)(\det B)$.

c. $\det A^{\mathrm{T}}=\det A$.

d. 若 A 是三角形矩阵，那么 $\det A$ 是 A 主对角线元素的乘积.

e. 对 A 作行替换不改变其行列式值. 作一次行交换使其行列式值符号改变一次. 数乘一行后，行列式值等于用此数乘原来的行列式值.

特征方程

利用定理 3 (a)，我们可以通过行列式来判断矩阵 $A-\lambda I$ 是否可逆. 数值方程 $\det(A-\lambda I)=0$ 称为 A 的**特征方程**，例 1 验证了以下结论.

数 λ 是 $n\times n$ 矩阵 A 的特征值的充要条件是 λ 是特征方程 $\det(A-\lambda I)=0$ 的根.

例 3 求 $A=\begin{bmatrix}5&-2&6&-1\\0&3&-8&0\\0&0&5&4\\0&0&0&1\end{bmatrix}$ 的特征方程.

解 写出 $A-\lambda I$ 并利用定理 3（d）：

$$\det(A-\lambda I)=\det\begin{bmatrix}5-\lambda & -2 & 6 & -1\\ 0 & 3-\lambda & -8 & 0\\ 0 & 0 & 5-\lambda & 4\\ 0 & 0 & 0 & 1-\lambda\end{bmatrix}$$
$$=(5-\lambda)(3-\lambda)(5-\lambda)(1-\lambda)$$

特征方程是

$$(5-\lambda)^2(3-\lambda)(1-\lambda)=0$$

或

$$(\lambda-5)^2(\lambda-3)(\lambda-1)=0$$

展开乘积，特征方程也可以写为

$$\lambda^4-14\lambda^3+68\lambda^2-130\lambda+75=0$$ ■

例 1 和例 3 的 $\det(A-\lambda I)$ 是关于 λ 的多项式. 可以看出，如果 A 是 $n\times n$ 矩阵，那么 $\det(A-\lambda I)$ 是 n 次多项式，称为 A 的**特征多项式**.

在例 3 中，由于因子 $(\lambda-5)$ 在特征多项式中出现了 2 次，故称特征值 5 有重数 2. 一般地，把特征值 λ 作为特征方程根的重数称为是 λ 的（**代数**）**重数**.

例 4 某 6×6 矩阵的特征多项式是 $\lambda^6-4\lambda^5-12\lambda^4$，求特征值及重数.

解 把多项式分解因式

$$\lambda^6-4\lambda^5-12\lambda^4=\lambda^4(\lambda^2-4\lambda-12)=\lambda^4(\lambda-6)(\lambda+2)$$

特征值是 0（重数为 4），6（重数为 1）和−2（重数为 1）. ■

我们同样可以说例 4 所给矩阵的特征值是 0，0，0，0，6 和−2，特征值按其重数重复计数.

因为 $n\times n$ 矩阵的特征方程包含有一个 n 次多项式，如果算上重根，并允许有复根，则特征方程恰好有 n 个根. 复根称为复特征值，将在 5.5 节讨论之. 目前，我们只考虑实特征值.

特征方程具有十分重要的理论意义. 但在实际计算时，任何大于 2×2 矩阵的特征值应该用计算机去求，除非矩阵是三角形或有其他特殊性质. 虽然 3×3 矩阵的特征多项式容易用手工算出，但对其因式分解却不是容易的事（除非是精心选择的矩阵）. 见本节末的数值计算注解.

相似性

下列定理说明了特征多项式的一个用途，并为某些近似计算特征值的迭代方法提供理论基础. 假如 A 和 B 是 $n\times n$ 矩阵，如果存在可逆矩阵 P，使得 $P^{-1}AP=B$，或等价地 $A=PBP^{-1}$，则称 A **相似于** B. 记 $Q=P^{-1}$，则有 $Q^{-1}BQ=A$，即 B 也相似于 A，故我们简单说 A 和 B 是**相似**的. 把 A 变成 $P^{-1}AP$ 的变换称为**相似变换**.

定理 4 若 $n\times n$ 矩阵 A 和 B 是相似的，那么它们有相同的特征多项式，从而有相同的特征值（和相同的重数）.

证 由 A 与 B 相似，有 $B=P^{-1}AP$，那么

$$B-\lambda I=P^{-1}AP-\lambda P^{-1}P=P^{-1}(AP-\lambda P)=P^{-1}(A-\lambda I)P$$

利用定理 3 的性质（b），有

$$\begin{aligned}\det(B-\lambda I) &= \det[P^{-1}(A-\lambda I)P] \\ &= \det(P^{-1})\cdot\det(A-\lambda I)\cdot\det(P)\end{aligned} \quad (2)$$

因为 $\det(P^{-1})\cdot\det(P)=\det(P^{-1}P)=\det I=1$，由（2）得 $\det(B-\lambda I)=\det(A-\lambda I)$ ■

警告 1. 矩阵$\begin{bmatrix}2 & 1\\0 & 2\end{bmatrix}$和$\begin{bmatrix}1 & 0\\0 & 2\end{bmatrix}$即使有相同的特征值也不相似.

2. 相似性与行等价不是一回事.（假如 A 行等价于 B，则存在可逆矩阵 E，使得 $1B=EA$）对矩阵作行变换通常会改变矩阵的特征值.

应用到动力系统

如本章引例中提到的，特征值与特征向量是我们剖析动力系统的离散演变的关键点.

例 5 设 $A=\begin{bmatrix}0.95 & 0.03\\0.05 & 0.97\end{bmatrix}$，分析由 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k\,(k=0,1,2,\cdots)$，$\boldsymbol{x}_0=\begin{bmatrix}0.6\\0.4\end{bmatrix}$所确定的动力系统的长期发展趋势.

解 第 1 步求 A 的特征值，并找出每个特征空间的基. A 的特征方程是

$$\begin{aligned}0 &= \det\begin{bmatrix}0.95-\lambda & 0.03\\0.05 & 0.97-\lambda\end{bmatrix}=(0.95-\lambda)(0.97-\lambda)-(0.03)(0.05)\\ &= \lambda^2-1.92\lambda+0.92\end{aligned}$$

由二次方程的求根公式

$$\begin{aligned}\lambda &= \frac{1.92\pm\sqrt{(1.92)^2-4(0.92)}}{2}=\frac{1.92\pm\sqrt{0.0064}}{2}\\ &= \frac{1.92\pm0.08}{2}=1 \text{ 或 } 0.92\end{aligned}$$

容易验证对应 $\lambda=1$ 和 $\lambda=0.92$ 的特征向量分别是 $\boldsymbol{v}_1=\begin{bmatrix}3\\5\end{bmatrix}$ 和 $\boldsymbol{v}_2=\begin{bmatrix}1\\-1\end{bmatrix}$的倍数.

下一步把给定的 $\boldsymbol{x}_0$ 表示为 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的线性组合，显然 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是 $\mathbb{R}^2$ 的基，因此存在系数 c_1 和 c_2，使得

$$\boldsymbol{x}_0=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2=[\boldsymbol{v}_1\ \ \boldsymbol{v}_2]\begin{bmatrix}c_1\\c_2\end{bmatrix} \quad (3)$$

事实上

$$\begin{aligned}\begin{bmatrix}c_1\\c_2\end{bmatrix} &= [\boldsymbol{v}_1\ \ \boldsymbol{v}_2]^{-1}\boldsymbol{x}_0=\begin{bmatrix}3 & 1\\5 & -1\end{bmatrix}^{-1}\begin{bmatrix}0.60\\0.40\end{bmatrix}\\ &= \frac{1}{-8}\begin{bmatrix}-1 & -1\\-5 & 3\end{bmatrix}\begin{bmatrix}0.60\\0.40\end{bmatrix}=\begin{bmatrix}0.125\\0.225\end{bmatrix}\end{aligned} \quad (4)$$

因为，式（3）的 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 是 A 的特征向量，故 $A\boldsymbol{v}_1=\boldsymbol{v}_1, A\boldsymbol{v}_2=(0.92)\boldsymbol{v}_2$，容易算出每个 $\boldsymbol{x}_k$：

$$\begin{aligned}\boldsymbol{x}_1=A\boldsymbol{x}_0 &= c_1A\boldsymbol{v}_1+c_2A\boldsymbol{v}_2 && \text{利用 } \boldsymbol{x}\mapsto A\boldsymbol{x} \text{ 的线性性质}\\ &= c_1\boldsymbol{v}_1+c_2(0.92)\boldsymbol{v}_2 && \boldsymbol{v}_1 \text{ 和 } \boldsymbol{v}_2 \text{ 是特征向量}\end{aligned}$$

$$\begin{aligned}\boldsymbol{x}_2 = A\boldsymbol{x}_1 &= c_1A\boldsymbol{v}_1 + c_2(0.92)A\boldsymbol{v}_2\\ &= c_1\boldsymbol{v}_1 + c_2(0.92)^2\boldsymbol{v}_2\end{aligned}$$

继续下去，有

$$\boldsymbol{x}_k = c_1\boldsymbol{v}_1 + c_2(0.92)^k\boldsymbol{v}_2 \quad (k = 0,1,2,\cdots)$$

把式（4）的 c_1 和 c_2 代入上式，得

$$\boldsymbol{x}_k = 0.125\begin{bmatrix}3\\5\end{bmatrix} + 0.225(0.92)^k\begin{bmatrix}1\\-1\end{bmatrix} \ (k = 0,1,2,\cdots) \tag{5}$$

$\boldsymbol{x}_k$ 的显式公式（5）就是差分方程 $\boldsymbol{x}_{k+1} = A\boldsymbol{x}_k$ 的解.

$$当\ k \to \infty\ 时，\ (0.92)^k \to 0, \boldsymbol{x}_k \to \begin{bmatrix}0.375\\0.625\end{bmatrix} = 0.125\boldsymbol{v}_1.$$ ■

把例 5 的计算结果应用于 4.9 节的马尔可夫链很有意思. 读过 4.9 节的读者可能注意到上面例 5 的矩阵 A 就是 4.9 节的移民矩阵 M，$\boldsymbol{x}_0$ 是城市和郊区的初始人口分布，$\boldsymbol{x}_k$ 表示 k 年后的人口分布.

4.9 节的定理 18 说明，对于 A 这样的矩阵，序列 $\boldsymbol{x}_k$ 趋向于一个稳态向量. 现在我们知道了原因所在，至少对移民矩阵是成立的. 这里稳态向量是 $0.125\boldsymbol{v}_1$，是特征向量 $\boldsymbol{v}_1$ 的倍数，计算 $\boldsymbol{x}_k$ 的公式（5）正好显示了为什么 $\boldsymbol{x}_k \to 0.125\boldsymbol{v}_1$.

数值计算的注解

1. 像 Mathematica 和 Maple 这样的计算机软件能利用符号计算求出一个中等规模矩阵的特征多项式. 但对于 $n \geqslant 5$ 的一般 $n \times n$ 矩阵，没有公式或有限算法来求解其特征方程.

2. 最佳的求特征值的数值方法完全避开特征多项式. 事实上，MATLAB 通过先求出矩阵 A 的特征值 $\lambda_1,\cdots,\lambda_n$，然后通过展开 $(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$ 的积来得到 A 的特征多项式.

3. 有几个基于定理 4 的估计矩阵 A 特征值的常用算法.非常有效的 QR 算法将在习题中讨论. 当 $A = A^{\mathrm{T}}$ 时，可使用称为雅可比的另一种方法来计算形如

$$A_1 = A \text{ 和 } A_{k+1} = P_k^{-1}A_kP_k \quad (k = 1,2,\cdots)$$

的矩阵序列.序列中的每个矩阵都相似于 A，因此有与 A 相同的特征值. 当 k 增大时，A_{k+1} 的非对角线元素趋于零，而主对角线上的元素近似于 A 的特征值.

4. 其他估计特征值的方法在 5.8 节讨论.

练习题

求 $A = \begin{bmatrix}1 & -4\\4 & 2\end{bmatrix}$ 的特征方程和特征值.

习题 5.2

求习题 1~8 矩阵的特征多项式和特征值.

1. $\begin{bmatrix}2 & 7\\7 & 2\end{bmatrix}$

2. $\begin{bmatrix}-4 & -1\\6 & 1\end{bmatrix}$

3. $\begin{bmatrix}-4 & 2\\6 & 7\end{bmatrix}$

4. $\begin{bmatrix}8 & 2\\3 & 3\end{bmatrix}$

5. $\begin{bmatrix}8 & 4\\4 & 8\end{bmatrix}$

6. $\begin{bmatrix}9 & 2\\2 & 5\end{bmatrix}$

7. $\begin{bmatrix}5 & 3\\-4 & 4\end{bmatrix}$

8. $\begin{bmatrix}-4 & 3\\2 & 1\end{bmatrix}$

习题 9~14 要用到 3.1 节的方法. 利用余子式展开或在 3.1 节习题 15~18 前给出的 3×3 行列式的特殊公式求矩阵的特征多项式.（提示：只用行变换求 3×3 矩阵的特征多项式不简便，因为涉及变量λ.）

9. $\begin{bmatrix} 4 & 0 & -1 \\ 0 & 4 & -1 \\ 1 & 0 & 2 \end{bmatrix}$　　10. $\begin{bmatrix} 3 & 1 & 1 \\ 0 & 5 & 0 \\ -2 & 0 & 7 \end{bmatrix}$

11. $\begin{bmatrix} 3 & 0 & 0 \\ 2 & 1 & 4 \\ 1 & 0 & 4 \end{bmatrix}$　　12. $\begin{bmatrix} -1 & 0 & 2 \\ 3 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$

13. $\begin{bmatrix} 6 & -2 & 0 \\ -2 & 9 & 0 \\ 5 & 8 & 3 \end{bmatrix}$　　14. $\begin{bmatrix} 4 & 0 & -1 \\ -1 & 0 & 4 \\ 0 & 2 & 3 \end{bmatrix}$

写出习题 15~17 中矩阵的特征值，特征值按其重数重复写出.

15. $\begin{bmatrix} 5 & 5 & 0 & 2 \\ 0 & 2 & -3 & 6 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 0 & 5 \end{bmatrix}$　　16. $\begin{bmatrix} 3 & 0 & 0 & 0 \\ 6 & 2 & 0 & 0 \\ 0 & 3 & 6 & 0 \\ 2 & 3 & 3 & -5 \end{bmatrix}$

17. $\begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ -5 & 1 & 0 & 0 & 0 \\ 3 & 8 & 0 & 0 & 0 \\ 0 & -7 & 2 & 1 & 0 \\ -4 & 1 & 9 & -2 & 3 \end{bmatrix}$

18. 能够证明特征值λ的代数重数大于或等于其特征空间的维数. 求下面矩阵 A 中的 h，使$\lambda=4$的特征空间是 2 维的.

$$A=\begin{bmatrix} 4 & 2 & 3 & 3 \\ 0 & 2 & h & 3 \\ 0 & 0 & 4 & 14 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

19. 设 A 是 $n\times n$ 矩阵，并假设 A 有 n 个实特征值 $\lambda_1,\cdots,\lambda_n$，特征值按其重数重复，因此

$$\det(A-\lambda I)=(\lambda_1-\lambda)(\lambda_2-\lambda)\cdots(\lambda_n-\lambda)$$

请解释为什么 $\det(A)=\lambda_1\cdot\lambda_2\cdots\lambda_n$（此结果对有复特征值的方阵仍然成立）.

20. 利用行列式的性质证明 A 和 A^{T} 有相同的特征多项式.

在习题 21 和习题 22 中，A 和 B 是 $n\times n$ 矩阵，判断每一命题的真假，并验证答案.

21. a. A 的行列式值等于其对角线上元素的乘积.
    b. 对 A 作初等行变换不会改变 A 的行列式值.
    c. $(\det A)(\det B)=\det AB$.
    d. 若 $\lambda+5$ 是 A 的特征多项式的因子，则 5 是 A 的特征值.

22. a. A 为 3×3 矩阵，其列为 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$，则 $\det A$ 等于由 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 生成的平行六面体的体积.
    b. $\det A^{\mathrm{T}}=(-1)\det A$.
    c. A 的特征方程的根 r 的重数称为特征值 r 的代数重数.
    d. 对 A 作行替换不会改变其特征值.

一个广泛用来估计一般矩阵 A 的特征值的方法是 QR 算法. 在适当的条件下，该算法产生一个矩阵序列，序列中的矩阵全部相似于 A，矩阵几乎是上三角的，并且主对角线上的元素近似于 A 的特征值. 主要思想是把 A（或另一个相似于 A 的矩阵）分解为 $A=Q_1R_1$，这里 $Q_1^{\mathrm{T}}=Q_1^{-1}$，而 R_1 是上三角矩阵. 将 Q_1 与 R_1 交换后形成 $A_1=R_1Q_1$，A_1 又被分解为 $A_1=Q_2R_2$，然后令 $A_2=R_2Q_2$，一直做下去. 由习题 23 的结论知，$A,A_1,\cdots$ 是相似的.

23. 证明：假如 $A=QR$，Q 可逆，则 A 相似于 $A_1=RQ$.

24. 证明：若 A 和 B 相似，则 $\det A=\det B$.

25. 设 $A=\begin{bmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{bmatrix},\boldsymbol{v}_1=\begin{bmatrix} 3/7 \\ 4/7 \end{bmatrix},\boldsymbol{x}_0=\begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$（注意：$A$ 是 4.9 节例 5 中研究的随机矩阵）.
    a. 求 $\mathbb{R}^2$ 的一个基，基由 $\boldsymbol{v}_1$ 和 A 的另一个特征向量组成.
    b. 证明 $\boldsymbol{x}_0$ 可以写成 $\boldsymbol{x}_0=\boldsymbol{v}_1+c\boldsymbol{v}_2$.
    c. 定义 $\boldsymbol{x}_k=A^k\boldsymbol{x}_0,k=1,2,\cdots$，证明当 k 增大时，$\boldsymbol{x}_k\to\boldsymbol{v}_1$.

26. 设 $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$，用本节给出的计算行列式的公式（1）证明 $\det A=ad-bc$，考虑 $a\neq 0$ 和 $a=0$

两种情况.

27. 设 $A=\begin{bmatrix}0.5&0.2&0.3\\0.3&0.8&0.3\\0.2&0&0.4\end{bmatrix},\boldsymbol{v}_1=\begin{bmatrix}0.3\\0.6\\0.1\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}1\\-3\\2\end{bmatrix},$

$\boldsymbol{v}_3=\begin{bmatrix}-1\\0\\1\end{bmatrix},\boldsymbol{w}=\begin{bmatrix}1\\1\\1\end{bmatrix}.$

a. 证明 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 是 A 的特征向量.（注意：A 是 4.9 节例 3 中的随机矩阵.）

b. 设 $\boldsymbol{x}_0$ 是 $\mathbb{R}^3$ 中的任一向量，其所有分量非负，且分量和为 1（在 4.9 节，$\boldsymbol{x}_0$ 被称为概率向量）. 解释为什么存在常数 c_1,c_2,c_3 使得 $\boldsymbol{x}_0=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+c_3\boldsymbol{v}_3$. 计算 $w^{\mathrm{T}}\boldsymbol{x}_0$，并推断 $c_1=1$.

c. 对（b）中的 $\boldsymbol{x}_0$，定义 $\boldsymbol{x}_k=A^k\boldsymbol{x}_0,k=1,2,\cdots$, 证明当 k 增大时，$\boldsymbol{x}_k\to\boldsymbol{v}_1$.

28. [M]构造一个随机的 4×4 整数值矩阵 A，并验证 A 和 A^{T} 有相同的特征多项式（相同的特征值和相同的特征值重数）. A 和 A^{T} 有相同的特征向量吗？对 5×5 矩阵作同样的分析，写出矩阵和计算结果.

29. [M]构造一个随机 4×4 整数值矩阵 A.

a. 不用倍乘将 A 化简为阶梯形矩阵 U，用（例 2 前的）（1）式中的 U 计算 $\det A$.（假如 A 是奇异的，重新构造一个随机矩阵再做.）

b. 计算 A 的特征值和特征值的乘积（尽可能精确）.

c. 写出矩阵 A，精确到三位小数，写出 U 的主元和 A 的特征值，用矩阵程序计算 $\det A$，并将结果与（a）和（b）得到的乘积进行比较.

30. [M] 设 $A=\begin{bmatrix}-6&28&21\\4&-15&-12\\-8&a&25\end{bmatrix}$，对 a 取集合 {32, 31.9,31.8, 32.1, 32.2} 中的每一个值，计算 A 的特征多项式和特征值，并画出相应的特征多项式 $p(t)=\det(A-tI)$ 在 $0\leqslant t\leqslant 3$ 的图. 有可能的话，在同一个坐标系中画出所有的图形，能否从图中看出特征值怎样随 a 的变化而变化？

练习题答案

特征方程是

$$0=\det(A-\lambda I)=\det\begin{bmatrix}1-\lambda&-4\\4&2-\lambda\end{bmatrix}$$
$$=(1-\lambda)\ (2-\lambda)-(-4)\ (4)=\lambda^2-3\lambda+18$$

解方程，求得

$$\lambda=\frac{3\pm\sqrt{(-3)^2-4(18)}}{2}=\frac{3\pm\sqrt{-63}}{2}$$

很明显，特征方程没有实数解，故没有实的特征值. 这样就不存在 $\mathbb{R}^2$ 的非零向量 $\boldsymbol{v}$ 和实数 λ 使得 $A\boldsymbol{v}=\lambda\boldsymbol{v}$ 成立.

5.3 对角化

在很多情况下，从分解式 $A=PDP^{-1}$，我们能够了解到有关矩阵 A 的特征值和特征向量的信息. 本节，我们利用分解式在 k 较大时能快速计算 A^k，这是线性代数中某些应用的基本思想. 稍后，在 5.6 和 5.7 节，分解式被用来分析（解耦）动力系统.

下面的例题表明对角矩阵的幂很容易计算.

例 1 若 $D=\begin{bmatrix}5&0\\0&3\end{bmatrix}$, 则 $D^2=\begin{bmatrix}5&0\\0&3\end{bmatrix}\begin{bmatrix}5&0\\0&3\end{bmatrix}=\begin{bmatrix}5^2&0\\0&3^2\end{bmatrix}$

$$D^3=DD^2=\begin{bmatrix}5&0\\0&3\end{bmatrix}\begin{bmatrix}5^2&0\\0&3^2\end{bmatrix}=\begin{bmatrix}5^3&0\\0&3^3\end{bmatrix}$$

一般有

$$\text{对 } k\geqslant 1,\ D^k=\begin{bmatrix}5^k&0\\0&3^k\end{bmatrix}$$ ■

若 $A=PDP^{-1}$，其中 P 为可逆矩阵，D 为对角矩阵，那么 A^k 的计算也很简单，见下例.

例 2 设 $A=\begin{bmatrix}7&2\\-4&1\end{bmatrix}$，给定 $A=PDP^{-1}$，其中 $P=\begin{bmatrix}1&1\\-1&-2\end{bmatrix}$, $D=\begin{bmatrix}5&0\\0&3\end{bmatrix}$，计算 A^k.

解 由 2×2 矩阵的逆矩阵的标准公式得

$$P^{-1}=\begin{bmatrix}2&1\\-1&-1\end{bmatrix}$$

然后，结合矩阵乘法，得

$$A^2=(PDP^{-1})(PDP^{-1})=PD\underbrace{(P^{-1}P)}_{I}DP^{-1}=PDDP^{-1}$$

$$=PD^2P^{-1}=\begin{bmatrix}1&1\\-1&-2\end{bmatrix}\begin{bmatrix}5^2&0\\0&3^2\end{bmatrix}\begin{bmatrix}2&1\\-1&-1\end{bmatrix}$$

同理，

$$A^3=(PDP^{-1})A^2=\underbrace{(PDP^{-1})P}_{I}D^2P^{-1}=PDD^2P^{-1}=PD^3P^{-1}$$

一般对 $k\geqslant 1$，有

$$A^k=PD^kP^{-1}=\begin{bmatrix}1&1\\-1&-2\end{bmatrix}\begin{bmatrix}5^k&0\\0&3^k\end{bmatrix}\begin{bmatrix}2&1\\-1&-1\end{bmatrix}$$

$$=\begin{bmatrix}2\cdot 5^k-3^k & 5^k-3^k\\ 2\cdot 3^k-2\cdot 5^k & 2\cdot 3^k-5^k\end{bmatrix}$$ ■

如果方阵 A 相似于对角矩阵，即存在可逆矩阵 P 和对角矩阵 D，有 $A=PDP^{-1}$，则称 A **可对角化**.下一个定理给出了可对角化矩阵的特征，并告诉我们如何去建立合适的分解式.

定理 5 （对角化定理）

$n\times n$ 矩阵 A 可对角化的充分必要条件是 A 有 n 个线性无关的特征向量.

事实上，$A=PDP^{-1}$，D 为对角矩阵的充分必要条件是 P 的列向量是 A 的 n 个线性无关的特征向量. 此时，D 的主对角线上的元素分别是 A 的对应于 P 中特征向量的特征值.

换句话说，A 可对角化的充分必要条件是有足够的特征向量形成 $\mathbb{R}^n$ 的基，我们称这样的基为**特征向量基**.

证　首先看到，若 P 是列为 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$ 的任一 $n\times n$ 矩阵，D 是对角线元素为 $\lambda_1,\cdots,\lambda_n$ 的对角矩阵，那么

$$AP = A[\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \cdots \quad \boldsymbol{v}_n] = [A\boldsymbol{v}_1 \quad A\boldsymbol{v}_2 \quad \cdots \quad A\boldsymbol{v}_n] \tag{1}$$

而

$$PD = P\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = [\lambda_1\boldsymbol{v}_1 \quad \lambda_2\boldsymbol{v}_2 \quad \cdots \quad \lambda_n\boldsymbol{v}_n] \tag{2}$$

现假设 A 可对角化且 $A=PDP^{-1}$，用 P 右乘等式两边，则有 $AP=PD$. 此时，由（1）和（2）得

$$[A\boldsymbol{v}_1 \quad A\boldsymbol{v}_2 \quad \cdots \quad A\boldsymbol{v}_n] = [\lambda_1\boldsymbol{v}_1 \quad \lambda_2\boldsymbol{v}_2 \quad \cdots \quad \lambda_n\boldsymbol{v}_n] \tag{3}$$

由列相等，有

$$A\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}_1,\ A\boldsymbol{v}_2 = \lambda_2\boldsymbol{v}_2,\cdots,A\boldsymbol{v}_n = \lambda_n\boldsymbol{v}_n \tag{4}$$

因为 P 可逆，故 P 的列 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$ 必定线性无关. 同样，因为这些 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$ 非零，（4）表示 $\lambda_1,\cdots,\lambda_n$ 是特征值，$\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$ 是相应的特征向量. 这就证明了定理中第一、第二和随后的第三个命题的必要性.

最后，给定任意 n 个特征向量 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$，用它们作为矩阵 P 的列，并用相应的特征值来构造矩阵 D，由（1）~（3），等式 $AP=PD$ 成立而不需要特征向量有任何条件.若特征向量是线性无关的，则 P 是可逆的（由可逆矩阵定理），由 $AP=PD$ 可推出 $A=PDP^{-1}$. ■

矩阵的对角化

例 3　可能的话，将下面矩阵对角化

$$A=\begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

即求可逆矩阵 P 和对角矩阵 D，使得 $A=PDP^{-1}$.

解　对角化工作可分为 4 步来完成.

第 1 步　求出 A 的特征值. 在 5.2 节曾提到，在矩阵的规模大于 2×2 时，可借用计算机求特征值，为避免分心，本书将会提供这一步的内容. 现在的特征方程是一个 3 次多项式，可分解为：

$$\begin{aligned} 0 = \det(A-\lambda I) &= -\lambda^3 - 3\lambda^2 + 4 \\ &= -(\lambda-1)(\lambda+2)^2 \end{aligned}$$

特征值是 $\lambda=1$ 和 $\lambda=-2$.

第 2 步　求 A 的 3 个线性无关的特征向量. 因为 A 是 3×3 的，故需要 3 个向量. 这一步很关键，如果找不到这 3 个向量，那么由定理 5，A 就不能对角化. 用 5.1 节的方法可求出每一特征空间的基：

$$对于\ \lambda=1\ 的基是：\ \boldsymbol{v}_1=\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

对于 $\lambda=-2$ 的基是：$v_2=\begin{bmatrix}-1\\1\\0\end{bmatrix}$ 和 $v_3=\begin{bmatrix}-1\\0\\1\end{bmatrix}$

你可以验证 $\{v_1,v_2,v_3\}$ 是线性无关的.

第 3 步　用第 2 步得到的向量构造矩阵 P. 向量的次序不重要，用第 2 步选择的次序，形成

$$p=[v_1\ \ v_2\ \ v_3]=\begin{bmatrix}1&-1&-1\\-1&1&0\\1&0&1\end{bmatrix}$$

第 4 步　用对应的特征值构造矩阵 D. 构造 D 时，特征值的次序必须和矩阵 P 选择的特征向量的次序一致. 对应 $\lambda=-2$ 的特征向量有 2 个，特征值 $\lambda=-2$ 要出现 2 次：

$$D=\begin{bmatrix}1&0&0\\0&-2&0\\0&0&-2\end{bmatrix}$$

验证所找的 P 和 D 是否正确是一个好的习惯，为避免计算 P^{-1}，可简单验证 $AP=PD$，这等价于当 P 可逆时 $A=PDP^{-1}$. （但必须确认 P 是可逆的!）我们计算

$$AP=\begin{bmatrix}1&3&3\\-3&-5&-3\\3&3&1\end{bmatrix}\begin{bmatrix}1&-1&-1\\-1&1&0\\1&0&1\end{bmatrix}=\begin{bmatrix}1&2&2\\-1&-2&0\\1&0&-2\end{bmatrix}$$

$$PD=\begin{bmatrix}1&-1&-1\\-1&1&0\\1&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&-2&0\\0&0&-2\end{bmatrix}=\begin{bmatrix}1&2&2\\-1&-2&0\\1&0&-2\end{bmatrix}$$ ■

例 4　可能的话，将下面的矩阵 A 对角化：

$$A=\begin{bmatrix}2&4&3\\-4&-6&-3\\3&3&1\end{bmatrix}$$

解　A 的特征方程与例 3 的完全一样：

$$0=\det(A-\lambda I)=-\lambda^3-3\lambda^2+4=-(\lambda-1)(\lambda+2)^2$$

特征值是 $\lambda=1$ 和 $\lambda=-2$. 但当我们找特征向量时，发现每一特征空间仅是一维的.

对于 $\lambda=1$ 的特征向量：$v_1=\begin{bmatrix}1\\-1\\1\end{bmatrix}$

对于 $\lambda=-2$ 的特征向量：$v_2=\begin{bmatrix}-1\\1\\0\end{bmatrix}$

没有其他的特征向量了，A 的每个特征向量都是 v_1 或 v_2 的倍数. 因此不可能利用 A 的特征向量

来构造出 $\mathbb{R}^3$ 的基. 由定理 5，A 不能对角化. ■

以下定理为矩阵可对角化提供了充分条件.

定理 6 有 n 个相异特征值的 $n\times n$ 矩阵可对角化.

证 设 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$ 是矩阵 A 对应于 n 个相异特征值的特征向量，由 5.1 节定理 2，$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$ 是线性无关的，因此，由定理 5，A 可对角化. ■

不过，$n\times n$ 矩阵并不需要有 n 个相异特征值才可对角化. 例 3 的 3×3 矩阵尽管只有 2 个相异的特征值，但它是可对角化的.

例 5 确定下列矩阵能否对角化：

$$A=\begin{bmatrix}5 & -8 & 1\\ 0 & 0 & 7\\ 0 & 0 & -2\end{bmatrix}$$

解 很容易判断矩阵 A 可对角化. 因为 A 是 3×3 的三角矩阵，其特征值显然是 5，0 和−2，有三个相异的特征值，故 A 可对角化. ■

特征值不是都相异的矩阵

如果 $n\times n$ 矩阵 A 有 n 个相异的特征值及相应的特征向量 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$，如果记 $P=[\boldsymbol{v}_1 \quad \cdots \quad \boldsymbol{v}_n]$，那么由定理 2，$P$ 的列是线性无关的，自然 P 是可逆的. 当 A 可对角化，但 A 相异的特征值的个数少于 n 时，我们仍可以用以下定理给出的方法来构造可逆矩阵 P[⊖].

定理 7 设 A 是 $n\times n$ 矩阵，其相异的特征值是 $\lambda_1,\ldots,\lambda_p$.

a. 对于 $1\leqslant k\leqslant p$，λ_k 的特征空间的维数小于或等于 λ_k 的代数重数.

b. 矩阵 A 可对角化的充分必要条件是所有不同特征空间的维数之和为 n. 即每个 λ_k 的特征空间的维数等于 λ_k 的代数重数.

c. 若 A 可对角化，$\mathcal{B}_k$ 是对应于 λ_k 的特征空间的基，那么，集合 $\mathcal{B}_1,\cdots,\mathcal{B}_p$ 中所有向量的集合是 $\mathbb{R}^n$ 的特征向量基.

例 6 可能的话，将下列矩阵对角化.

$$A=\begin{bmatrix}5 & 0 & 0 & 0\\ 0 & 5 & 0 & 0\\ 1 & 4 & -3 & 0\\ -1 & -2 & 0 & -3\end{bmatrix}$$

解 A 是三角矩阵，特征值是 $\lambda=5$ 和 $\lambda=-3$，重数都是 2. 利用 5.1 节的方法，我们求出每个特征空间的基.

⊖ 定理 7 的证明有点长但不难，可看 S.Friedberg, A. Insel, and L.Spence，*Linear Algebra*，4th ed.（Englewood Cliffs,NJ: Prentice-Hall, 2002），5.2 节.

$$\text{对应}\lambda=5\text{的特征向量：}\quad v_1=\begin{bmatrix}-8\\4\\1\\0\end{bmatrix}\text{和}v_2=\begin{bmatrix}-16\\4\\0\\1\end{bmatrix}$$

$$\text{对应}\lambda=-3\text{的特征向量：}\quad v_3=\begin{bmatrix}0\\0\\1\\0\end{bmatrix}\text{和}v_4=\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

由定理 7，向量集合 $\{v_1,\dots,v_4\}$ 是线性无关的，故 $P=[v_1 \ \cdots \ v_4]$ 是可逆的，且有 $A=PDP^{-1}$，其中

$$P=\begin{bmatrix}-8&-16&0&0\\4&4&0&0\\1&0&1&0\\0&1&0&1\end{bmatrix},\ D=\begin{bmatrix}5&0&0&0\\0&5&0&0\\0&0&-3&0\\0&0&0&-3\end{bmatrix}$$ ■

练习题

1. $A=\begin{bmatrix}4&-3\\2&-1\end{bmatrix}$，计算 A^8.
2. 设 $A=\begin{bmatrix}-3&12\\-2&7\end{bmatrix}, v_1=\begin{bmatrix}3\\1\end{bmatrix}, v_2=\begin{bmatrix}2\\1\end{bmatrix}$，若已知 v_1 和 v_2 是 A 的特征向量，将 A 对角化.
3. 设 4×4 矩阵 A 的特征值是 5，3 和−2，并且已知对应 $\lambda=3$ 的特征空间是 2 维的，你能否判断 A 是可对角化的？

习题 5.3

在习题 1 和习题 2 中，设 $A=PDP^{-1}$，计算 A^4.

1. $P=\begin{bmatrix}5&7\\2&3\end{bmatrix}, D=\begin{bmatrix}2&0\\0&1\end{bmatrix}$
2. $P=\begin{bmatrix}1&2\\2&3\end{bmatrix}, D=\begin{bmatrix}1&0\\0&3\end{bmatrix}$

在习题 3 和习题 4 中，利用分解式 $A=PDP^{-1}$，计算 A^k，k 为正整数.

3. $\begin{bmatrix}a&0\\2(a-b)&b\end{bmatrix}=\begin{bmatrix}1&0\\2&1\end{bmatrix}\begin{bmatrix}a&0\\0&b\end{bmatrix}\begin{bmatrix}1&0\\-2&1\end{bmatrix}$
4. $\begin{bmatrix}1&-6\\2&-6\end{bmatrix}=\begin{bmatrix}3&2\\2&-1\end{bmatrix}\begin{bmatrix}-3&0\\0&-2\end{bmatrix}\begin{bmatrix}-1&2\\-2&3\end{bmatrix}$

在习题 5 和习题 6 中，矩阵 A 被分解为 PDP^{-1}，利用对角化定理求 A 的特征值和每个特征空间的基.

5. $\begin{bmatrix}2&-1&-1\\1&4&1\\-1&-1&2\end{bmatrix}$
$=\begin{bmatrix}1&-1&0\\-1&1&-1\\0&-1&1\end{bmatrix}\begin{bmatrix}3&0&0\\0&2&0\\0&0&3\end{bmatrix}\begin{bmatrix}0&-1&-1\\-1&-1&-1\\-1&-1&0\end{bmatrix}$
6. $\begin{bmatrix}3&0&0\\-3&4&9\\0&0&3\end{bmatrix}$
$=\begin{bmatrix}3&0&-1\\0&1&-3\\1&0&0\end{bmatrix}\begin{bmatrix}3&0&0\\0&4&0\\0&0&3\end{bmatrix}\begin{bmatrix}0&0&1\\-3&1&9\\-1&0&3\end{bmatrix}$

可能的话，将习题 7~20 的矩阵对角化，习题 11~16 和习题 18 的特征值在矩阵下方给出.

7. $\begin{bmatrix} 1 & 0 \\ 6 & -1 \end{bmatrix}$

8. $\begin{bmatrix} 3 & 2 \\ 0 & 3 \end{bmatrix}$

9. $\begin{bmatrix} 2 & -1 \\ 1 & 4 \end{bmatrix}$

10. $\begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}$

11. $\begin{bmatrix} 0 & 1 & 1 \\ 2 & 1 & 2 \\ 3 & 3 & 2 \end{bmatrix}$ $\lambda = -1, 5$

12. $\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$ $\lambda = 2, 5$

13. $\begin{bmatrix} 2 & 2 & -1 \\ 1 & 3 & -1 \\ -1 & -2 & 2 \end{bmatrix}$ $\lambda = 1, 5$

14. $\begin{bmatrix} 2 & 0 & -2 \\ 1 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix}$ $\lambda = 2, 3$

15. $\begin{bmatrix} 0 & -1 & -1 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$ $\lambda = 0, 1$

16. $\begin{bmatrix} 1 & 2 & -3 \\ 2 & 5 & -2 \\ 1 & 3 & 1 \end{bmatrix}$ $\lambda = 0$

17. $\begin{bmatrix} 2 & 0 & 0 \\ 2 & 2 & 0 \\ 2 & 2 & 2 \end{bmatrix}$

18. $\begin{bmatrix} 2 & -2 & -2 \\ 3 & -3 & -2 \\ 2 & -2 & -2 \end{bmatrix}$ $\lambda = -2, -1, 0$

19. $\begin{bmatrix} 5 & -3 & 0 & 9 \\ 0 & 3 & 1 & -2 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

20. $\begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 3 \end{bmatrix}$

在习题 21 和习题 22 中，A, B, P 和 D 是 $n \times n$ 矩阵，标出每个命题的真假，证明你的答案（在做这些习题之前先仔细研读定理 5 和定理 6 及本节的例子）.

21. a. 若存在矩阵 D 和可逆矩阵 P 使 $A=PDP^{-1}$ 成立，则 A 可对角化.
    b. 若 $\mathbb{R}^n$ 有 A 的特征向量基，则 A 可对角化.
    c. A 可对角化的充分必要条件是 A 有 n 个特征值，算上重数.
    d. 若 A 可对角化，则 A 是可逆的.
22. a. 若 A 有 n 个特征向量，则 A 可对角化.
    b. 若 A 可对角化，则 A 有 n 个相异的特征值.
    c. 若 $AP = PD$ ，D 为对角矩阵，那么 P 的非零列必定是 A 的特征向量.
    d. 若 A 可逆，则 A 可对角化.
23. 5×5 矩阵 A 有 2 个特征值，一个特征空间是 3 维的，另一个特征空间是 2 维的，A 可对角化否？为什么？
24. 3×3 矩阵 A 有 2 个特征值，每个特征空间是 1 维的，A 可对角化否？为什么？
25. 4×4 矩阵 A 有 3 个特征值，有一个特征空间为 1 维，其他特征空间之一是 2 维的，A 是否可对角化？说明理由.
26. 7×7 矩阵 A 有 3 个特征值，一个特征空间是 2 维的，其他特征空间之一是 3 维的，A 是否不可对角化？说明理由.
27. 证明：若 A 可对角化且可逆，则 A^{-1} 亦可对角化.
28. 证明：若 A 有 n 个线性无关的特征向量，则 A^T 也有 n 个线性无关的特征向量.（提示：用对角化定理.）
29. 分解式 $A = PDP^{-1}$ 不是唯一的，用例 2 的矩阵 A 来证实这一事实.若 $D_1 = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}$，利用例 2 的结果求矩阵 P_1，使得 $A = P_1D_1P_1^{-1}$.
30. 对例 2 的 A 和 D，求不等于 P 的可逆矩阵 P_2，使得 $A = P_2DP_2^{-1}$.
31. 构造一个非零的 2×2 可逆但不可对角化的矩阵.
32. 构造一个不是对角阵的 2×2 可对角化但不可逆的矩阵.

[M]将习题 33~36 的矩阵对角化，先用矩阵程序的特征值命令求特征值，然后用 5.1 节的方法找出特征空间的基.

33. $\begin{bmatrix} 9 & -4 & -2 & -4 \\ -56 & 32 & -28 & 44 \\ -14 & -14 & 6 & -14 \\ 42 & -33 & 21 & -45 \end{bmatrix}$

34. $\begin{bmatrix} 4 & -9 & -7 & 8 & 2 \\ -7 & -9 & 0 & 7 & 14 \\ 5 & 10 & 5 & -5 & -10 \\ -2 & 3 & 7 & 0 & 4 \\ -3 & -13 & 7 & 10 & 11 \end{bmatrix}$

35. $\begin{bmatrix} 13 & -12 & 9 & -15 & 9 \\ 6 & -5 & 9 & -15 & 9 \\ 6 & -12 & -5 & 6 & 9 \\ 6 & -12 & 9 & -8 & 9 \\ -6 & 12 & 12 & -6 & -2 \end{bmatrix}$

36. $\begin{bmatrix} 24 & -6 & 2 & 6 & 2 \\ 72 & 51 & 9 & -99 & 9 \\ 0 & -63 & 15 & 63 & 63 \\ 72 & 15 & 9 & -63 & 9 \\ 0 & 63 & 21 & -63 & -27 \end{bmatrix}$

练习题答案

1. $\det(A-\lambda I)=\lambda^2-3\lambda+2=(\lambda-2)(\lambda-1)$，特征值是 2 和 1，对应的特征向量是 $\boldsymbol{v}_1=\begin{bmatrix}3\\2\end{bmatrix}$ 和 $\boldsymbol{v}_2=\begin{bmatrix}1\\1\end{bmatrix}$，然后构造 $P=\begin{bmatrix}3 & 1\\2 & 1\end{bmatrix}, D=\begin{bmatrix}2 & 0\\0 & 1\end{bmatrix}, P^{-1}=\begin{bmatrix}1 & -1\\-2 & 3\end{bmatrix}$. 由 $A=PDP^{-1}$，得

$$A^8=PD^8P^{-1}=\begin{bmatrix}3 & 1\\2 & 1\end{bmatrix}\begin{bmatrix}2^8 & 0\\0 & 1^8\end{bmatrix}\begin{bmatrix}1 & -1\\-2 & 3\end{bmatrix}=\begin{bmatrix}3 & 1\\2 & 1\end{bmatrix}\begin{bmatrix}256 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}1 & -1\\-2 & 3\end{bmatrix}=\begin{bmatrix}766 & -765\\510 & -509\end{bmatrix}$$

2. 计算 $A\boldsymbol{v}_1=\begin{bmatrix}-3 & 12\\-2 & 7\end{bmatrix}\begin{bmatrix}3\\1\end{bmatrix}=\begin{bmatrix}3\\1\end{bmatrix}=1\cdot\boldsymbol{v}_1$，$A\boldsymbol{v}_2=\begin{bmatrix}-3 & 12\\-2 & 7\end{bmatrix}\begin{bmatrix}2\\1\end{bmatrix}=\begin{bmatrix}6\\3\end{bmatrix}=3\cdot\boldsymbol{v}_2$

   显然，$\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 分别是特征值 1 和 3 对应的特征向量，因此 $A=PDP^{-1}$，这里

$$P=\begin{bmatrix}3 & 2\\1 & 1\end{bmatrix} \text{和} D=\begin{bmatrix}1 & 0\\0 & 3\end{bmatrix}$$

3. A 可对角化，对应 $\lambda=3$ 的特征空间有一个基 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$. 此外，对应 $\lambda=5$ 和 $\lambda=-2$ 各取一个特征向量，记为 $\boldsymbol{v}_3$ 和 $\boldsymbol{v}_4$，则 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_4\}$ 是线性无关的. 由定理 7，A 可对角化. 因为向量全属于 $\mathbb{R}^4$，$\mathbb{R}^4$ 的基包含的向量个数不超过 4，故对应于 $\lambda=5$ 和 $\lambda=-2$ 的特征空间都是 1 维的.

5.4 特征向量与线性变换

在本节，我们将矩阵分解 $A=PDP^{-1}$ 理解为线性变换. 我们还将看到变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 实质上是简单的映射 $\boldsymbol{u}\to D\boldsymbol{u}$. 即使 D 不是对角矩阵时，对矩阵 A 和 D 仍有相似的解释.

在 1.9 节曾讲到，任意一个从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性变换 T 可通过左乘矩阵 A 来实现，矩阵 A 称为 T 的标准矩阵. 现在，我们对两个有限维向量空间之间的线性变换也作同样的描述.

线性变换的矩阵

设 V 是 n 维向量空间，W 是 m 维向量空间，T 是 V 到 W 的线性变换. 为了把 T 与矩阵相联系，我们指定 $\mathcal{B}$ 和 $\mathcal{C}$ 分别是 V 和 W 的基.

若 $\boldsymbol{x}\in V$，坐标向量 $[\boldsymbol{x}]_{\mathcal{B}}\in\mathbb{R}^n$，$\boldsymbol{x}$ 的像 $T(\boldsymbol{x})$ 的坐标向量 $[T(\boldsymbol{x})]_{\mathcal{C}}\in\mathbb{R}^m$，见图 5-6.

容易找出 $[\boldsymbol{x}]_{\mathcal{B}}$ 和 $[T(\boldsymbol{x})]_{\mathcal{C}}$ 之间的关系.设 V 的基 $\mathcal{B}$ 是 $\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$. 若 $\boldsymbol{x}=r_1\boldsymbol{b}_1+\cdots+r_n\boldsymbol{b}_n$，则

$$[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}r_1\\ \vdots\\ r_n\end{bmatrix}$$

因为 T 是线性的，故

$$T(\boldsymbol{x})=T(r_1\boldsymbol{b}_1+\cdots+r_n\boldsymbol{b}_n)=r_1T(\boldsymbol{b}_1)+\cdots+r_nT(\boldsymbol{b}_n) \tag{1}$$

因为从 W 到 $\mathbb{R}^m$ 的坐标映射是线性的（4.4 节定理 8），等式（1）可推出

$$[T(\boldsymbol{x})]_{\mathcal{C}}=r_1[T(\boldsymbol{b}_1)]_{\mathcal{C}}+\cdots+r_n[T(\boldsymbol{b}_n)]_{\mathcal{C}} \tag{2}$$

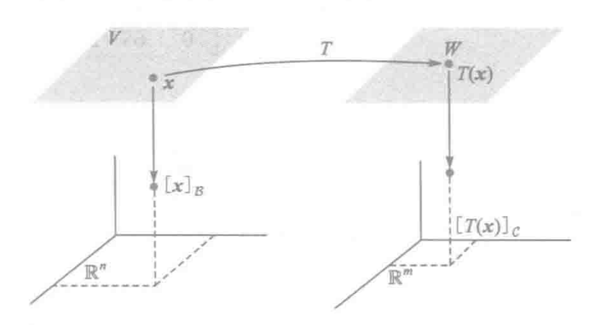


图 5-6 V 到 W 的线性变换

因为这些$\mathcal{C}$–坐标向量都属于 $\mathbb{R}^m$，向量等式（2）可以写为矩阵等式

$$[T(\boldsymbol{x})]_{\mathcal{C}}=M[\boldsymbol{x}]_{\mathcal{B}} \tag{3}$$

其中

$$M=[[T(\boldsymbol{b}_1)]_{\mathcal{C}}\quad[T(\boldsymbol{b}_2)]_{\mathcal{C}}\quad\cdots\quad[T(\boldsymbol{b}_n)]_{\mathcal{C}}] \tag{4}$$

矩阵 M 是 T 的矩阵表示，称为 **T 相对于基$\mathcal{B}$和$\mathcal{C}$的矩阵**. 见图 5-7.

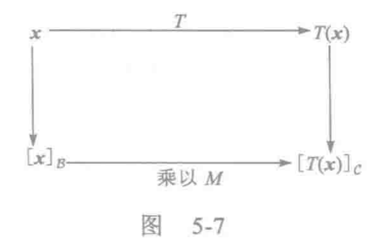


图 5-7

等式（3）表明，就坐标向量而言，T 对 $\boldsymbol{x}$ 的作用相当于用矩阵 M 左乘 $\boldsymbol{x}$.

例 1 设 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$ 是 V 的基，$\mathcal{C}=\{\boldsymbol{c}_1,\boldsymbol{c}_2,\boldsymbol{c}_3\}$ 是 W 的基. T 是 $V\to W$ 的线性变换，满足性质：

$$T(\boldsymbol{b}_1)=3\boldsymbol{c}_1-2\boldsymbol{c}_2+5\boldsymbol{c}_3,\ T(\boldsymbol{b}_2)=4\boldsymbol{c}_1+7\boldsymbol{c}_2-\boldsymbol{c}_3$$

求 T 相对于基$\mathcal{B}$和$\mathcal{C}$的矩阵 M .

解 $\boldsymbol{b}_1$ 和 $\boldsymbol{b}_2$ 的像的$\mathcal{C}$–坐标向量是

$$[T(\boldsymbol{b}_1)]_{\mathcal{C}}=\begin{bmatrix}3\\-2\\5\end{bmatrix},\qquad [T(\boldsymbol{b}_2)]_{\mathcal{C}}=\begin{bmatrix}4\\7\\-1\end{bmatrix}$$

因此

$$M=\begin{bmatrix}3&4\\-2&7\\5&-1\end{bmatrix}$$

■

如果$\mathcal{B}$和$\mathcal{C}$是同一空间 V 的基，T 是线性变换$T(\boldsymbol{x})=\boldsymbol{x},\boldsymbol{x}\in V$，那么（4）中的矩阵 M 正好是坐标变换矩阵（见 4.7 节）.

V 到 V 的线性变换

当$W=V,\mathcal{C}=\mathcal{B}$时，（4）中的 M 称为 **T 相对于 $\mathcal{B}$ 的矩阵**，或简称为 **T 的 $\mathcal{B}$–矩阵**，记为$[T]_{\mathcal{B}}$，见图 5-8.

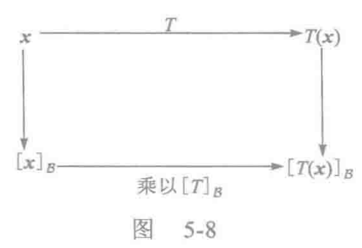


图 5-8

$V\to V$ 的线性变换 T 的 $\mathcal{B}$–矩阵对所有 V 中的 $\boldsymbol{x}$，有

$$[T(\boldsymbol{x})]_{\mathcal{B}}=[T]_{\mathcal{B}}[\boldsymbol{x}]_{\mathcal{B}} \tag{5}$$

例 2 $\mathbb{P}_2\to\mathbb{P}_2$ 的映射 T：$T(a_0+a_1t+a_2t^2)=a_1+2a_2t$ 是线性变换.（学过微积分的学生知道 T 是微分算子.）

a. 当基$\mathcal{B}=\{1,t,t^2\}$时，求 T 的 $\mathcal{B}$–矩阵.

b. 对 $\mathbb{P}_2$ 中的每个 $\boldsymbol{p}$，验证$[T(\boldsymbol{p})]_{\mathcal{B}}=[T]_{\mathcal{B}}[\boldsymbol{p}]_{\mathcal{B}}$.

解 a. 计算基向量的像

$$T(1)=0 \qquad \text{零多项式}$$
$$T(t)=1 \qquad \text{值恒为 1 的多项式}$$
$$T(t^2)=2t$$

然后写出$T(1),T(t),T(t^2)$的$\mathcal{B}$–坐标（本例通过观察可以得到），把它们放在一起组成T的矩阵：

$$[T(1)]_{\mathcal{B}}=\begin{bmatrix}0\\0\\0\end{bmatrix},\ [T(t)]_{\mathcal{B}}=\begin{bmatrix}1\\0\\0\end{bmatrix},\ [T(t^2)]_{\mathcal{B}}=\begin{bmatrix}0\\2\\0\end{bmatrix}$$

$$[T]_{\mathcal{B}}=\begin{bmatrix}0&1&0\\0&0&2\\0&0&0\end{bmatrix}$$

b. 对一般的多项式 $\boldsymbol{p}(t)=a_0+a_1t+a_2t^2$，我们有

$$[T(\boldsymbol{p})]_{\mathcal{B}}=[a_1+2a_2t]_{\mathcal{B}}=\begin{bmatrix}a_1\\2a_2\\0\end{bmatrix}$$

$$=\begin{bmatrix}0&1&0\\0&0&2\\0&0&0\end{bmatrix}\begin{bmatrix}a_0\\a_1\\a_2\end{bmatrix}=[T]_{\mathcal{B}}[\boldsymbol{p}]_{\mathcal{B}}$$

见图 5-9. ■

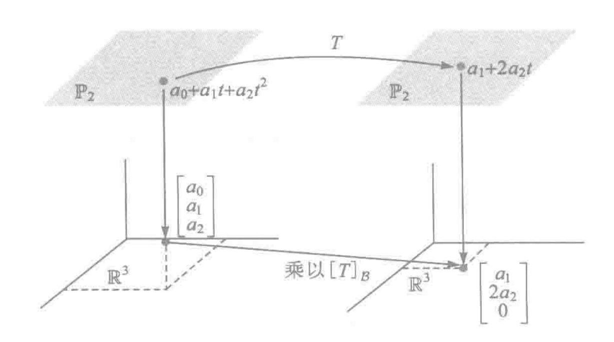


图 5-9 线性变换的矩阵表示

$\mathbb{R}^n$ 上的线性变换

在涉及$\mathbb{R}^n$ 的应用问题中，线性变换首先表现为一个矩阵变换 $\boldsymbol{x} \mapsto A\boldsymbol{x}$. 假设 A 是可对角化的，那么存在由 A 的特征向量组成的$\mathbb{R}^n$ 的基. 此时，下面的定理 8 表明 T 的$\mathcal{B}$–矩阵是对角矩阵，这样，把 A 对角化相当于找到变换 $\boldsymbol{x} \mapsto A\boldsymbol{x}$ 的对角矩阵表示.

定理 8 （对角矩阵表示）

设 $A=PDP^{-1}$，其中 D 为 $n\times n$ 对角矩阵，若$\mathbb{R}^n$ 的基$\mathcal{B}$由 P 的列向量组成，那么 D 是变换 $\boldsymbol{x} \mapsto A\boldsymbol{x}$ 的$\mathcal{B}$–矩阵.

证 记 P 的列向量为 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n$，则有 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}, P=[\boldsymbol{b}_1 \cdots \boldsymbol{b}_n]$. 此时，$P$ 是 4.4 节中讨论过的坐标变换矩阵 $P_{\mathcal{B}}$，其中

$$P[\boldsymbol{x}]_{\mathcal{B}} = \boldsymbol{x} \text{ 和 } [\boldsymbol{x}]_{\mathcal{B}} = P^{-1}\boldsymbol{x}$$

若 $\boldsymbol{x} \in \mathbb{R}^n, T(\boldsymbol{x}) = A\boldsymbol{x}$，则

$$\begin{aligned}
[T]_{\mathcal{B}} &= [[T(\boldsymbol{b}_1)]_{\mathcal{B}} \cdots [T(\boldsymbol{b}_n)]_{\mathcal{B}}] && \text{由 } [T]_{\mathcal{B}} \text{ 的定义} \\
&= [[A\boldsymbol{b}_1]_{\mathcal{B}} \cdots [A\boldsymbol{b}_n]_{\mathcal{B}}] && \text{由 } T(\boldsymbol{x}) = A\boldsymbol{x} \\
&= [P^{-1}A\boldsymbol{b}_1 \cdots P^{-1}A\boldsymbol{b}_n] && \text{坐标变换} \\
&= P^{-1}A[\boldsymbol{b}_1 \cdots \boldsymbol{b}_n] && \text{矩阵乘法} \\
&= P^{-1}AP
\end{aligned} \tag{6}$$

由于 $A = PDP^{-1}$，我们有 $[T]_{\mathcal{B}} = P^{-1}AP = D$. ■

例 3 设 $A=\begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$, $\mathbb{R}^2 \to \mathbb{R}^2$ 的变换 $T: T(\boldsymbol{x}) = A\boldsymbol{x}$. 求 $\mathbb{R}^2$ 的一个基 $\mathcal{B}$，使得 T 的 $\mathcal{B}$–矩阵是对角矩阵.

解 由 5.3 节的例 2，我们知道 $A = PDP^{-1}$，其中

$$P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix},\ D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

P 的列向量，记为 $\boldsymbol{b}_1$ 和 $\boldsymbol{b}_2$，是 A 的特征向量，由定理 8，若取 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$，则 D 是 T 的 $\mathcal{B}-$矩阵. 映射 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 和 $\boldsymbol{u}\mapsto D\boldsymbol{u}$ 描述的是相对于不同基的同一个线性变换. ■

矩阵表示的相似性

定理 8 的证明并没有用到 D 是对角矩阵这一事实. 因此，若 A 相似于 C，即有 $A=PCP^{-1}$，如果 $\mathcal{B}$ 由 P 的列向量组成，则 C 是变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的 $\mathcal{B}-$矩阵. 分解 $A=PCP^{-1}$ 如图 5-10 所示.

图 5-10　两个矩阵表示的相似性 $A=PCP^{-1}$

相反，若 $\mathbb{R}^n\to\mathbb{R}^n$ 的变换 $T:T(\boldsymbol{x})=A\boldsymbol{x}$，而 $\mathcal{B}$ 是 $\mathbb{R}^n$ 的任意一个基，那么 T 的 $\mathcal{B}-$矩阵相似于 A. 其实，在定理 8 的证明的计算中已经证明若 P 是以 $\mathcal{B}$ 的向量作为列构成的矩阵，那么 $[T]_{\mathcal{B}}=P^{-1}AP$. 因此，所有相似于 A 的矩阵的集合与变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的所有矩阵表示的集合是同一集合.

例 4　设 $A=\begin{bmatrix}4&-9\\4&8\end{bmatrix}$，$\boldsymbol{b}_1=\begin{bmatrix}3\\2\end{bmatrix}$，$\boldsymbol{b}_2=\begin{bmatrix}2\\1\end{bmatrix}$，$A$ 的特征多项式是 $(\lambda+2)^2$，但特征值–2 的特征空间只是一维的，因此 A 是不可以对角化的. 但是，有一组基 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$ 能够使得变换 $\boldsymbol{x}\to A\boldsymbol{x}$ 的 $\mathcal{B}-$矩阵是三角矩阵，称为 A 的约当形.[⊖] 求 A 的 $\mathcal{B}-$矩阵.

解　如果 $P=[\boldsymbol{b}_1\quad \boldsymbol{b}_2]$，则 $\mathcal{B}-$矩阵是 $P^{-1}AP$，计算

$$AP=\begin{bmatrix}4&-9\\4&-8\end{bmatrix}\begin{bmatrix}3&2\\2&1\end{bmatrix}=\begin{bmatrix}-6&-1\\-4&0\end{bmatrix}$$

$$P^{-1}AP=\begin{bmatrix}-1&2\\2&-3\end{bmatrix}\begin{bmatrix}-6&-1\\-4&0\end{bmatrix}=\begin{bmatrix}-2&1\\0&-2\end{bmatrix}$$

注意 A 的特征值在对角线上. ■

数值计算的注解　计算 $\mathcal{B}-$矩阵 $P^{-1}AP$ 的一个有效方法是先计算 AP，然后用行变换将增广矩阵 $[P\quad AP]$ 化为 $[I\quad P^{-1}AP]$，这样就不需要单独计算 P^{-1} 了，见 2.2 节的习题 15.

练习题

1. 假设 T 是 $\mathbb{P}_2$ 到 $\mathbb{P}_2$ 的线性变换，相对于基 $\mathcal{B}=\{1,t,t^2\}$ 的矩阵是 $[T]_{\mathcal{B}}=\begin{bmatrix}3&4&0\\0&5&-1\\1&-2&7\end{bmatrix}$，求 $T(a_0+a_1t+a_2t^2)$.

2. 设 A,B,C 是 $n\times n$ 矩阵，我们已经知道若 A 相似于 B，那么 B 亦相似于 A，这个性质和下面的结论说明

⊖ 每个方阵都与一个若尔当型矩阵相似. 可以导出若尔当型的基包含有 A 的特征向量和所谓的“推广的特征向量”. 见 B.Nobel 和 J.W.Daniel 合写的 *Applied Linear Algebra*,3rd ed.（Englewood Cliffs, NJ: Prentice-Hall,1988）第 9 章.

"相似"是等价关系.(行等价是等价关系的另一个例子.)证明(a)和(b).

a. A 相似于 A.

b. 若 A 相似于 B,B 相似于 C,那么 A 相似于 C.

习题 5.4

1. 设 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3\}$ 和 $\mathcal{D}=\{\boldsymbol{d}_1,\boldsymbol{d}_2\}$ 分别是向量空间 V 和 W 的基,$T:V\to W$ 是线性变换,满足:
$T(\boldsymbol{b}_1)=3\boldsymbol{d}_1-5\boldsymbol{d}_2, T(\boldsymbol{b}_2)=-\boldsymbol{d}_1+6\boldsymbol{d}_2, T(\boldsymbol{b}_3)=4\boldsymbol{d}_2$
求 T 相对于 $\mathcal{B}$ 和 $\mathcal{D}$ 的矩阵.

2. 设 $\mathcal{D}=\{\boldsymbol{d}_1,\boldsymbol{d}_2\}$ 和 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$ 分别是向量空间 V 和 W 的基,$T:V\to W$ 是线性变换,满足:
$$T(\boldsymbol{d}_1)=3\boldsymbol{b}_1-3\boldsymbol{b}_2,\ T(\boldsymbol{d}_2)=-2\boldsymbol{b}_1+5\boldsymbol{b}_2$$
求 T 相对于 $\mathcal{D}$ 和 $\mathcal{B}$ 的矩阵.

3. 设 $\mathcal{E}=\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}$ 是 $\mathbb{R}^3$ 的标准基,$\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3\}$ 是向量空间 V 的基,$T:\mathbb{R}^3\to V$ 是线性变换,
$T(x_1,x_2,x_3)=(2x_3-x_2)\boldsymbol{b}_1-(2x_2)\boldsymbol{b}_2+(x_1+3x_3)\boldsymbol{b}_3$
   a. 计算 $T(\boldsymbol{e}_1),T(\boldsymbol{e}_2),T(\boldsymbol{e}_3)$.
   b. 计算 $[T(\boldsymbol{e}_1)]_{\mathcal{B}},[T(\boldsymbol{e}_2)]_{\mathcal{B}},[T(\boldsymbol{e}_3)]_{\mathcal{B}}$.
   c. 求 T 相对于$\mathcal{E}$和 $\mathcal{B}$ 的矩阵.

4. 设 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3\}$ 是向量空间 V 的基,$T:V\to\mathbb{R}^2$ 是线性变换,满足:
$$T(x_1\boldsymbol{b}_1+x_2\boldsymbol{b}_2+x_3\boldsymbol{b}_3)=\begin{bmatrix}2x_1-3x_2+x_3\\-2x_2+5x_3\end{bmatrix}$$
求 T 相对于 $\mathcal{B}$ 和 $\mathbb{R}^2$ 的标准基的矩阵.

5. 设 $T:\mathbb{P}_2\to\mathbb{P}_3$ 是将多项式 $\boldsymbol{p}(t)$ 映射为多项式 $(t+3)\boldsymbol{p}(t)$ 的变换.
   a. 求 $\boldsymbol{p}(t)=3-2t+t^2$ 的像.
   b. 证明 T 是线性变换.
   c. 求 T 相对于基 $\{1,t,t^2\}$ 和 $\{1,t,t^2,t^3\}$ 的矩阵.

6. 设 $T:\mathbb{P}_2\to\mathbb{P}_4$ 是将多项式 $\boldsymbol{p}(t)$映射为 $\boldsymbol{p}(t)+2t^2\boldsymbol{p}(t)$ 的变换.
   a. 求 $\boldsymbol{p}(t)=3-2t+t^2$ 的像.
   b. 证明 T 是线性变换.
   c. 求 T 相对于基 $\{1,t,t^2\}$ 和 $\{1,t,t^2,t^3,t^4\}$ 的矩阵.

7. 假设映射 $T:\mathbb{P}_2\to\mathbb{P}_2$ 定义为
$T(a_0+a_1t+a_2t^2)=3a_0+(5a_0-2a_1)t+(4a_1+a_2)t^2$
并且是线性的,求 T 相对于基 $\mathcal{B}=\{1,t,t^2\}$ 的矩阵.

8. 设 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3\}$是向量空间 V 的基,T 是 $V\to V$ 的线性变换,其相对于 $\mathcal{B}$ 的矩阵是
$[T]_{\mathcal{B}}=\begin{bmatrix}0&0&1\\2&1&-2\\1&3&1\end{bmatrix}$,求 $T(4\boldsymbol{b}_1-3\boldsymbol{b}_2)$.

9. 定义 $\mathbb{P}_2\to\mathbb{R}^3$ 的变换 T 为 $T(\boldsymbol{p})=\begin{bmatrix}\boldsymbol{p}(-1)\\\boldsymbol{p}(0)\\\boldsymbol{p}(1)\end{bmatrix}$.
   a. 求 $\boldsymbol{p}(t)=5+3t$ 在变换 T 下的像.
   b. 证明 T 是线性变换.
   c. 求 T 相对于 $\mathbb{P}_2$ 的基 $\{1,t,t^2\}$ 和 $\mathbb{R}^3$ 标准基的矩阵.

10. 定义 $\mathbb{P}_3\to\mathbb{R}^4$ 的变换 T 为 $T(\boldsymbol{p})=\begin{bmatrix}\boldsymbol{p}(-2)\\\boldsymbol{p}(3)\\\boldsymbol{p}(1)\\\boldsymbol{p}(0)\end{bmatrix}$.
   a. 证明 T 是线性变换.
   b. 求 T 相对于 $\mathbb{P}_3$ 的基 $\{1,t,t^2,t^3\}$ 和 $\mathbb{R}^4$ 标准基的矩阵.

在习题 11 和习题 12 中,$\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$,求变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的 $\mathcal{B}-$矩阵.

11. $A=\begin{bmatrix}-4&-1\\6&1\end{bmatrix}$, $\boldsymbol{b}_1=\begin{bmatrix}-1\\2\end{bmatrix}$, $\boldsymbol{b}_2=\begin{bmatrix}-1\\1\end{bmatrix}$

12. $A=\begin{bmatrix}-6&-2\\4&0\end{bmatrix}$, $\boldsymbol{b}_1=\begin{bmatrix}0\\1\end{bmatrix}$, $\boldsymbol{b}_2=\begin{bmatrix}-1\\2\end{bmatrix}$

在习题 13~16 中,定义 $\mathbb{R}^2\to\mathbb{R}^2$ 的变换 T 为 $T(\boldsymbol{x})=A\boldsymbol{x}$,求 $\mathbb{R}^2$ 的基 $\mathcal{B}$,使得 $[T]_{\mathcal{B}}$ 为对角矩阵.

13. $A=\begin{bmatrix}0&1\\-3&4\end{bmatrix}$

14. $A=\begin{bmatrix}2&3\\3&2\end{bmatrix}$

15. $A=\begin{bmatrix}1&2\\3&-4\end{bmatrix}$

16. $A=\begin{bmatrix}4&2\\-1&5\end{bmatrix}$

17. 设 $A=\begin{bmatrix}4&1\\-1&2\end{bmatrix}$ 和 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}, \boldsymbol{b}_1=\begin{bmatrix}1\\-1\end{bmatrix}$,

$\boldsymbol{b}_2=\begin{bmatrix}-1\\2\end{bmatrix}$定义 $T:\mathbb{R}^2\to\mathbb{R}^2$ 为 $T(\boldsymbol{x})=A\boldsymbol{x}$.

a. 证明 $\boldsymbol{b}_1$ 是 A 的特征向量，但 A 不可对角化.

b. 求 T 的 $\mathcal{B}-$矩阵.

18. 定义 $T:\mathbb{R}^3\to\mathbb{R}^3$ 为 $T(\boldsymbol{x})=A\boldsymbol{x}$ ，其中 A 为 3×3 矩阵，特征值是 5，5 和–2. 是否存在 $\mathbb{R}^3$ 的基，使得 T 的 $\mathcal{B}-$矩阵是对角矩阵？试做讨论.

证明习题 19~24 的命题，题中矩阵为方阵.

19. 若 A 可逆且相似于 B ，则 B 可逆且 A^{-1} 相似于 B^{-1} .（提示：存在可逆矩阵 P ，使得 $P^{-1}AP=B$ ，解释 B 为什么可逆，然后找一个可逆矩阵 Q ，使得 $Q^{-1}A^{-1}Q=B^{-1}$.）

20. 若 A 相似于 B ，则 A^2 相似于 B^2 .

21. 若 B 相似于 A ， C 相似于 A ，则 B 相似于 C .

22. 若 A 可对角化，B 相似于 A ，则 B 亦可对角化.

23. 若 $B=P^{-1}AP,\boldsymbol{x}$ 是 A 对应特征值 λ 的特征向量，那么 $P^{-1}\boldsymbol{x}$ 是 B 对应 λ 的特征向量.

24. 若 A 与 B 是相似的，则它们有相同的秩.（提示：参考第 4 章补充习题 13 和 14.）

25. 方阵 A 主对角线元素之和称为 A 的**轨迹**，记为 tr A. 可以证明对任意的两个 $n\times n$ 矩阵 F 和 G ，有 $\mathrm{tr}(FG)=\mathrm{tr}(GF)$. 证明，若 A 与 B 相似，则 tr A=tr B.

26. 能够证明矩阵 A 的迹等于 A 的特征值之和，在 A 可对角化时验证这个结论.

27. 设 $V=\mathbb{R}^n,\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ 为 V 的基， $W=\mathbb{R}^n,\mathcal{E}$ 为 W 的标准基. $I:I(\boldsymbol{x})=\boldsymbol{x}$ 是 $\mathbb{R}^n\to\mathbb{R}^n$ 的恒等变换，求 I 相对于 $\mathcal{B}$ 和 $\mathcal{E}$ 的矩阵，这个矩阵在 4.4 节被称为什么矩阵？

28. 设 V 和 W 是同一向量空间，其基分别是 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ 和 $\mathcal{C}=\{\boldsymbol{c}_1,\cdots,\boldsymbol{c}_n\},I$ 是 $V\to W$ 的恒等变换，求 I 相对于 $\mathcal{B}$ 和 $\mathcal{C}$ 的矩阵.这个矩阵在 4.7 节被称为什么矩阵？

29. 设 V 是向量空间，基为 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ ，求 $V\to V$ 的恒等变换 I 的 $\mathcal{B}-$矩阵.

[M]对习题 30 和习题 31，求变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 相对基 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3\}$ 的 $\mathcal{B}-$矩阵.

30. $A=\begin{bmatrix}6&-2&-2\\3&1&-2\\2&-2&2\end{bmatrix},\boldsymbol{b}_1=\begin{bmatrix}1\\1\\1\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}2\\1\\3\end{bmatrix},$

$\boldsymbol{b}_3=\begin{bmatrix}-1\\-1\\0\end{bmatrix}$

31. $A=\begin{bmatrix}-7&-48&-16\\1&14&6\\-3&-45&-19\end{bmatrix},\boldsymbol{b}_1=\begin{bmatrix}-3\\1\\-3\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}-2\\1\\-3\end{bmatrix},$

$\boldsymbol{b}_3=\begin{bmatrix}3\\-1\\0\end{bmatrix}$

32. [M]变换 T 的标准矩阵为 A ，求 $\mathbb{R}^4$ 的一个基 $\mathcal{B}$ ，使得 $[T]_{\mathcal{B}}$ 为对角矩阵.

$$A=\begin{bmatrix}-6&-4&0&9\\-3&0&1&6\\-1&-2&1&0\\-4&-4&0&7\end{bmatrix}$$

练习题答案

1. 设 $\boldsymbol{p}(t)=a_0+a_1t+a_2t^2$ ，计算得

$$[T(\boldsymbol{p})]_{\mathcal{B}}=[T]_{\mathcal{B}}[\boldsymbol{p}]_{\mathcal{B}}=\begin{bmatrix}3&4&0\\0&5&-1\\1&-2&7\end{bmatrix}\begin{bmatrix}a_0\\a_1\\a_2\end{bmatrix}=\begin{bmatrix}3a_0+4a_1\\5a_1-a_2\\a_0-2a_1+7a_2\end{bmatrix}$$

故 $T(\boldsymbol{p})=(3a_0+4a_1)+(5a_1-a_2)t+(a_0-2a_1+7a_2)t^2$.

2. a. $A=(I)^{-1}AI$ ，所以 A 相似于 A .

b. 由假设，存在可逆矩阵 P 和 Q ，使得 $B=P^{-1}AP$ 和 $C=Q^{-1}BQ$ ，代入并利用逆运算性质，有 $C=Q^{-1}BQ=Q^{-1}(P^{-1}AP)Q=(PQ)^{-1}A(PQ)$ ，这一等式表明 A 相似于 C .

5.5 复特征值

$n\times n$ 矩阵的特征方程含有 n 次多项式，如果考虑复根，方程恰好有 n 个根，重根重复计算. 本节我们将看到假如矩阵 A 的特征方程有复根，那么这些复根将给我们提供有关 A 的关键信息. 关键的问题是让 A 作用于 n 维复空间 $\mathbb{C}^n$ 中的 n 元复数.㊀

我们对 $\mathbb{C}^n$ 感兴趣并不是为了把前几章的结果推广，尽管这可以开辟线性代数应用的新领域㊁. 然而，对复特征值的研究能使我们去揭示各种实际生活问题中出现的某些实矩阵中隐藏的信息. 这些问题包括很多蕴涵周期运动的实动力系统、振动或空间的某种旋转.

建立在 $\mathbb{R}^n$ 基础上的矩阵特征值、特征向量理论同样可以很好地应用到 $\mathbb{C}^n$. 因此，一个复数 λ 满足 $\det(A-\lambda I)=0$ 当且仅当在 $\mathbb{C}^n$ 中存在一个非零向量 $\boldsymbol{x}$，使得 $A\boldsymbol{x}=\lambda\boldsymbol{x}$. 我们称这样的 λ 是复特征值，$\boldsymbol{x}$ 是对应 λ 的（复）特征向量.

例 1 假设 $A=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$，那么 $\mathbb{R}^2$ 上的线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 将平面逆时针旋转 1/4 圈. A 的作用是周期性的，因为在旋转 4 次 1/4 圈后，向量又回到了它的初始位置，很明显没有非零向量被映射成自身的数倍，所以 A 在 $\mathbb{R}^2$ 中没有特征向量，因此也没有实的特征值. 实际上，A 的特征方程是

$$\lambda^2+1=0$$

只有复根 $\lambda=\mathrm{i}$ 和 $\lambda=-\mathrm{i}$. 但是，如果我们让 A 作用在 $\mathbb{C}^2$ 上，那么

$$\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix}=\begin{bmatrix}\mathrm{i}\\ 1\end{bmatrix}=\mathrm{i}\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix}$$

$$\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix}=\begin{bmatrix}-\mathrm{i}\\ 1\end{bmatrix}=-\mathrm{i}\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix}$$

因此 i 和 $-\mathrm{i}$ 是特征值，$\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix}$ 和 $\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix}$ 是对应的特征向量（在例 2 中讨论求复特征向量的方法）. ■

下一个例子的矩阵是本节的主要焦点所在.

例 2 设 $A=\begin{bmatrix}0.5 & -0.6\\ 0.75 & 1.1\end{bmatrix}$，求 A 的特征值及每个特征空间的基.

解 A 的特征方程是

$$0=\det\begin{bmatrix}0.5-\lambda & -0.6\\ 0.75 & 1.1-\lambda\end{bmatrix}=(0.5-\lambda)(1.1-\lambda)-(-0.6)(0.75)=\lambda^2-1.6\lambda+1$$

依求根公式，得 $\lambda=\frac{1}{2}\left[1.6\pm\sqrt{(-1.6)^2-4}\right]=0.8\pm0.6\mathrm{i}$. 对特征值 $\lambda=0.8-0.6\mathrm{i}$，构造

㊀ 参考附录 B 中对复数的简短讨论，有关实向量空间的矩阵代数和概念同样适用于复元素和复数的情形. 特别地，对于有复元素的 $m\times n$ 矩阵 A，$\boldsymbol{x},\boldsymbol{y}\in\mathbb{C}^n$ 及 $c,d\in\mathbb{C}$，有 $A(c\boldsymbol{x}+d\boldsymbol{y})=cA\boldsymbol{x}+dA\boldsymbol{y}$.

㊁ 线性代数的另一课程通常讨论这些主题，它们在电子工程中非常重要.

$$A-(0.8-0.6\mathrm{i})I=\begin{bmatrix}0.5 & -0.6\\0.75 & 1.1\end{bmatrix}-\begin{bmatrix}0.8-0.6\mathrm{i} & 0\\0 & 0.8-0.6\mathrm{i}\end{bmatrix}=\begin{bmatrix}-0.3+0.6\mathrm{i} & -0.6\\0.75 & 0.3+0.6\mathrm{i}\end{bmatrix} \quad (1)$$

由于有复数运算，手工做对增广矩阵行化简是相当慢的. 不过，细心观察就能找到简化问题的方法：因为 $0.8-0.6\mathrm{i}$ 是特征值，我们知道方程组

$$\begin{aligned}(-0.3+0.6\mathrm{i})x_1-\quad\quad 0.6x_2&=0\\0.75x_1+(0.3+0.6\mathrm{i})x_2&=0\end{aligned} \quad (2)$$

有非零解（这里的 x_1 和 x_2 可能是复数），因此，（2）中的 2 个方程确定的 x_1 与 x_2 之间的关系是同一关系，这样就可以通过其中的一个方程将某个变量用另一个来表示.㊀

由（2）的第 2 个方程得

$$0.75x_1=(-0.3-0.6\mathrm{i})x_2$$
$$x_1=(-0.4-0.8\mathrm{i})x_2$$

为去掉小数，取 $x_2=5$，有 $x_1=-2-4\mathrm{i}$. 对应 $\lambda=0.8-0.6\mathrm{i}$ 的特征空间的基是

$$\boldsymbol{v}_1=\begin{bmatrix}-2-4\mathrm{i}\\5\end{bmatrix}$$

对 $\lambda=0.8+0.6\mathrm{i}$ 做同样的计算，可求出特征向量

$$\boldsymbol{v}_2=\begin{bmatrix}-2+4\mathrm{i}\\5\end{bmatrix}$$

为验证结果是否正确，可以计算

$$A\boldsymbol{v}_2=\begin{bmatrix}0.5 & -0.6\\0.75 & 1.1\end{bmatrix}\begin{bmatrix}-2+4\mathrm{i}\\5\end{bmatrix}=\begin{bmatrix}-4+2\mathrm{i}\\4+3\mathrm{i}\end{bmatrix}=(0.8+0.6\mathrm{i})\boldsymbol{v}_2$$ ■

非常意外，例 2 的矩阵 A 确定的变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 实质上是一个旋转变换，画出一些适当的点后，这一事实会变得更明显.

例 3 与例 2 中的矩阵 A 相乘去影响点，通过以下方法可以看到这种影响. 先画出一个随机的初始点（如 $\boldsymbol{x}_0=(2,0)$），然后画出重复乘 A 后该点的后续图像. 即，画出

$$\boldsymbol{x}_1=A\boldsymbol{x}_0=\begin{bmatrix}0.5 & -0.6\\0.75 & 1.1\end{bmatrix}\begin{bmatrix}2\\0\end{bmatrix}=\begin{bmatrix}1.0\\1.5\end{bmatrix}$$

$$\boldsymbol{x}_2=A\boldsymbol{x}_1=\begin{bmatrix}0.5 & -0.6\\0.75 & 1.1\end{bmatrix}\begin{bmatrix}1.0\\1.5\end{bmatrix}=\begin{bmatrix}-0.4\\2.4\end{bmatrix}$$

$$\boldsymbol{x}_3=A\boldsymbol{x}_2$$

$$\cdots$$

在图 5-11 中，用粗点显示 $\boldsymbol{x}_0,\cdots,\boldsymbol{x}_8$，用细点显示的是 $\boldsymbol{x}_9,\cdots,\boldsymbol{x}_{100}$ 的位置，序列分布在椭圆形的轨道上.

㊀ 另一种理解是（1）中的矩阵是不可逆的，因此，它的行向量线性相关（类似 $\mathbb{C}^2$ 中的向量）故其中的一行是另一行的（复）数倍.

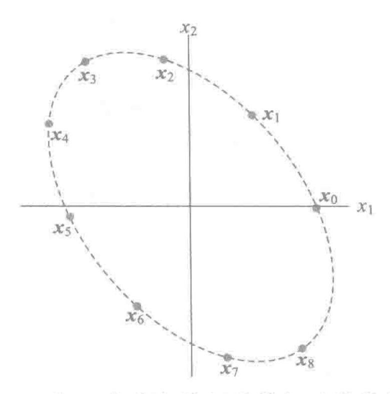


图 5-11 点 $\boldsymbol{x}_0$ 在有复特征值的矩阵的作用下的迭代 ■

当然，图 5-11 没有解释为什么会发生旋转，旋转的秘密隐藏在复特征向量的实部和虚部.

向量的实部和虚部

$\mathbb{C}^n$ 的复向量 x 的共轭向量 $\overline{x}$ 也是 $\mathbb{C}^n$ 的向量，它的分量是 $\boldsymbol{x}$ 对应分量的共轭复数，向量 $\mathrm{Re}\,\boldsymbol{x}$ 和 $\mathrm{Im}\,\boldsymbol{x}$ 称为复向量 $\boldsymbol{x}$ 的实部和虚部，分别由 $\boldsymbol{x}$ 的分量的实部和虚部组成.

例 4 假如 $x=\begin{bmatrix}3-\mathrm{i}\\ \mathrm{i}\\ 2+5\mathrm{i}\end{bmatrix}=\begin{bmatrix}3\\0\\2\end{bmatrix}+\mathrm{i}\begin{bmatrix}-1\\1\\5\end{bmatrix}$，那么

$$\mathrm{Re}\,\boldsymbol{x}=\begin{bmatrix}3\\0\\2\end{bmatrix},\ \mathrm{Im}\,\boldsymbol{x}=\begin{bmatrix}-1\\1\\5\end{bmatrix},\ \overline{\boldsymbol{x}}=\begin{bmatrix}3\\0\\2\end{bmatrix}-\mathrm{i}\begin{bmatrix}-1\\1\\5\end{bmatrix}=\begin{bmatrix}3+\mathrm{i}\\-\mathrm{i}\\2-5\mathrm{i}\end{bmatrix}$$ ■

假设 B 是可能有复元素的 $m\times n$ 矩阵，那么，以 B 的元素的共轭复数为元素的矩阵记为 $\overline{B}$. 复数的共轭运算性质对复矩阵代数亦成立：

$$\overline{r\boldsymbol{x}}=\overline{r}\ \overline{\boldsymbol{x}},\ \overline{B\boldsymbol{x}}=\overline{B}\ \overline{\boldsymbol{x}},\ \overline{BC}=\overline{B}\ \overline{C},\ \overline{rB}=\overline{r}\ \overline{B}$$

作用于 $\mathbb{C}^n$ 上的实矩阵的特征值和特征向量

设 A 为 $n\times n$ 的实矩阵，那么 $\overline{A\boldsymbol{x}}=\overline{A}\ \overline{\boldsymbol{x}}=A\overline{\boldsymbol{x}}$，假如 λ 是 A 的特征值，$\boldsymbol{x}$ 是对应 λ 的特征向量，那么

$$A\overline{\boldsymbol{x}}=\overline{A\boldsymbol{x}}=\overline{\lambda\boldsymbol{x}}=\overline{\lambda}\ \overline{\boldsymbol{x}}$$

故 $\overline{\lambda}$ 同样是 A 的特征值，而 $\overline{\boldsymbol{x}}$ 是对应的特征向量. 这表明，当 A 是实矩阵时，它的复特征值以共轭复数对出现（在这里或别的地方，我们用术语“复特征值”来表示形如 $\lambda=a+b\mathrm{i}, b\neq 0$ 的特征值）.

例 5 例 2 的实矩阵的特征值是共轭复数对 $0.8-0.6\mathrm{i}$ 和 $0.8+0.6\mathrm{i}$. 对应的特征向量同样是共轭复向量

$$\boldsymbol{v}_1=\begin{bmatrix}-2-4\mathrm{i}\\5\end{bmatrix}\text{和}\ \boldsymbol{v}_2=\begin{bmatrix}-2+4\mathrm{i}\\5\end{bmatrix}=\overline{\boldsymbol{v}}_1$$ ■

例 6 为计算有复特征值的 2×2 实矩阵提供了基本模式.

例 6 设 $C=\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$，其中 a,b 为实数且不都等于零. 那么 C 的特征值是 $\lambda=a\pm b\mathrm{i}$.（见本节末尾的练习题.）同样，假如 $r=|\lambda|=\sqrt{a^2+b^2}$，那么

$$C=r\begin{bmatrix} a/r & -b/r \\ b/r & a/r \end{bmatrix}=\begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

这里的 φ 是正 x 轴与 (0 , 0) 到 (a ,b) 射线的夹角. 看图 5-12 和附录 B，角 φ 称为 $\lambda=a\pm b\mathrm{i}$ 的辐角. 因此变换 $\boldsymbol{x}\mapsto C\boldsymbol{x}$ 可看作由旋转 φ 角度和倍乘 $|\lambda|$ 变换复合而成（见图 5-13）.

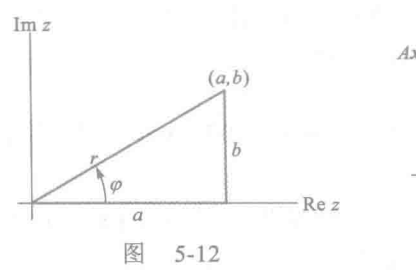


图 5-12

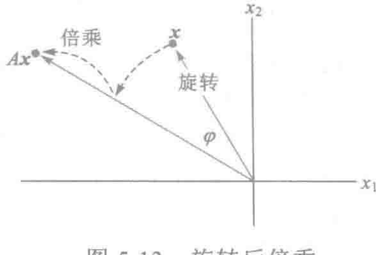


图 5-13 旋转后倍乘 ■

最后，我们准备讨论有复特征值的实矩阵中隐含的旋转.

例 7 与例 2 一样，$A=\begin{bmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{bmatrix}$，$\lambda=0.8-0.6\mathrm{i}$，$\boldsymbol{v}_1=\begin{bmatrix} -2-4\mathrm{i} \\ 5 \end{bmatrix}$. 同时，设 P 是 2×2 实矩阵，

$$P=\begin{bmatrix}\mathrm{Re}\,\boldsymbol{v}_1 & \mathrm{Im}\,\boldsymbol{v}_1\end{bmatrix}=\begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix}$$

并令

$$C=P^{-1}AP=\frac{1}{20}\begin{bmatrix} 0 & 4 \\ -5 & -2 \end{bmatrix}\begin{bmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{bmatrix}\begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix}=\begin{bmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{bmatrix}$$

由例 6，因为 $|\lambda|^2=(0.8)^2+(0.6)^2=1$，故 C 仅是旋转变换. 由 $C=P^{-1}AP$，得

$$A=PCP^{-1}=P\begin{bmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{bmatrix}P^{-1}$$

A “含有”旋转！矩阵 P 提供变量代换，如 $\boldsymbol{x}=P\boldsymbol{u}$. A 的作用相当于将 $\boldsymbol{x}$ 代换为 $\boldsymbol{u}$，再经过旋转，然后又代换回初始变量，见图 5-14. 旋转产生一个椭圆，如图 5-11 所示，而不是圆，因为由 P 的列确定的坐标系不是长方形的，在两个轴上没有相等的单位长.

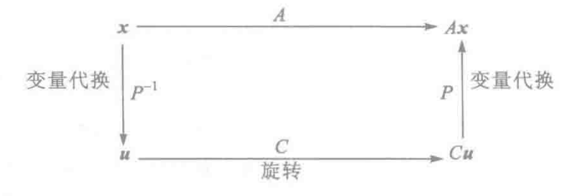


图 5-14 由复特征值引起的旋转 ■

下列定理表明例 7 的计算适用于任意有复特征值 λ 的 2×2 实矩阵. 证明利用了结论：若 A 是实矩阵，则 $A(\mathrm{Re}\,\boldsymbol{x})=\mathrm{Re}\,A\boldsymbol{x}$ 和 $A(\mathrm{Im}\,\boldsymbol{x})=\mathrm{Im}\,A\boldsymbol{x}$，若 $\boldsymbol{x}$ 是对应复特征值的特征向量，则 $\mathrm{Re}\,\boldsymbol{x}$ 和 $\mathrm{Im}\,\boldsymbol{x}$ 是线性无关的. （见习题 25 和 26.）这里省略细节.

定理 9 设 A 是 2×2 实矩阵，有复特征值 $\lambda=a-b\mathrm{i}(b\neq 0)$ 及对应的 $\mathbb{C}^2$ 中的复特征向量 $\boldsymbol{v}$，那么

$$A=PCP^{-1}\text{，其中 } P=\begin{bmatrix}\mathrm{Re}\,\boldsymbol{v} & \mathrm{Im}\,\boldsymbol{v}\end{bmatrix}, C=\begin{bmatrix}a & -b\\ b & a\end{bmatrix}$$

在更高维矩阵中亦存在例 7 中出现的现象. 例如，若 A 是有复特征值的 3×3 矩阵，那么在 $\mathbb{R}^3$ 中存在某个平面，A 对平面的作用是旋转（可能还结合倍乘），平面中的每个向量被旋转到该平面的另一点上，我们说平面在 A 的作用下是**不变**的.

例 8 矩阵 $A=\begin{bmatrix}0.8 & -0.6 & 0\\ 0.6 & 0.8 & 0\\ 0 & 0 & 1.07\end{bmatrix}$ 有特征值 $0.8\pm 0.6\mathrm{i}$ 和 1.07. x_1x_2 平面（第 3 坐标为 0）的任一向量 $\boldsymbol{w}_0$ 被 A 旋转到该平面的另一位置上.不在该平面的任一向量 $\boldsymbol{x}_0$ 的 $\boldsymbol{x}_3$ 坐标乘 1.07，图 5-15 显示了点 $\boldsymbol{w}_0=(2,0,0)$ 和 $\boldsymbol{x}_0=(2,0,1)$ 乘 A 后的迭代结果.

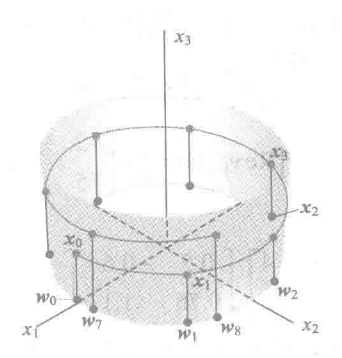


图 5-15 点 $\boldsymbol{w}_0$ 和 $\boldsymbol{x}_0$ 在有复特征值的 3×3 矩阵作用下的迭代 ■

练习题

证明：若 a 和 b 是实数，则 $A=\begin{bmatrix}a & -b\\ b & a\end{bmatrix}$ 的特征值是 $a\pm b\mathrm{i}$，对应的特征向量是 $\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix}$ 和 $\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix}$.

习题 5.5

让习题 1~6 的每个矩阵作用在 $\mathbb{C}^2$，求矩阵的特征值及对应 $\mathbb{C}^2$ 中特征空间的基.

1. $\begin{bmatrix}1 & -2\\ 1 & 3\end{bmatrix}$
2. $\begin{bmatrix}3 & -3\\ 3 & 3\end{bmatrix}$
3. $\begin{bmatrix}5 & 1\\ -8 & 1\end{bmatrix}$
4. $\begin{bmatrix}1 & -2\\ 1 & 3\end{bmatrix}$
5. $\begin{bmatrix}3 & 1\\ -2 & 5\end{bmatrix}$
6. $\begin{bmatrix}7 & -5\\ 1 & 3\end{bmatrix}$

在习题 7~12 中，用例 6 写出 A 的特征值. 在每一例中，变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 由旋转加倍乘复合而成. 求旋转的角度 φ（$-\pi<\varphi\leqslant\pi$）和倍乘因子 r.

7. $\begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}$　　8. $\begin{bmatrix} 3 & 3\sqrt{3} \\ -3\sqrt{3} & 3 \end{bmatrix}$

9. $\begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$　　10. $\begin{bmatrix} 0 & 0.5 \\ -0.5 & 0 \end{bmatrix}$

11. $\begin{bmatrix} -\sqrt{3} & 1 \\ -1 & -\sqrt{3} \end{bmatrix}$　　12. $\begin{bmatrix} 3 & -\sqrt{3} \\ \sqrt{3} & 3 \end{bmatrix}$

在习题 13~20 中，求可逆矩阵 P 和形如 $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ 的矩阵 C，把所给的矩阵表示为 $A=PCP^{-1}$.

13. $\begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}$　　14. $\begin{bmatrix} 3 & -3 \\ 1 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 0 & 5 \\ -2 & 2 \end{bmatrix}$　　16. $\begin{bmatrix} 4 & -2 \\ 1 & 6 \end{bmatrix}$

17. $\begin{bmatrix} -11 & -4 \\ 20 & 5 \end{bmatrix}$　　18. $\begin{bmatrix} 3 & -5 \\ 2 & 5 \end{bmatrix}$

19. $\begin{bmatrix} 1.52 & -0.7 \\ 0.56 & 0.4 \end{bmatrix}$　　20. $\begin{bmatrix} -3 & -8 \\ 4 & 5 \end{bmatrix}$

21. 在例 2 中，解（2）中的第 1 个方程，用 x_1 表示 x_2，并由此求得 A 的特征向量 $\boldsymbol{y}=\begin{bmatrix} 2 \\ -1+2\mathrm{i} \end{bmatrix}$. 证明 $\boldsymbol{y}$ 是例 2 中 $\boldsymbol{v}_1$ 的（复）数倍.

22. 设 A 是 $n\times n$ 的复（或实）矩阵，$\boldsymbol{x}$ 是复特征值对应 $\mathbb{C}^n$ 中的复特征向量. 证明：对任意非零的复数 u，向量 $u\boldsymbol{x}$ 是 A 的特征向量.

第 7 章将把重点放在具有性质 $A^{\mathrm{T}}=A$ 的矩阵 A 上，习题 23 和 24 证明了这种矩阵的特征值必定是实数.

23. 设 $n\times n$ 实矩阵 A 有性质 $A^{\mathrm{T}}=A$，$\boldsymbol{x}$ 是 $\mathbb{C}^n$ 中的向量，令 $q=\overline{\boldsymbol{x}}^{\mathrm{T}}A\boldsymbol{x}$. 下面的等式通过验证 $\overline{q}=q$ 证明了 q 是实数. 给出每一步的理由.
$$\overline{q}=\overline{\overline{\boldsymbol{x}}^{\mathrm{T}}A\boldsymbol{x}}\underset{(\mathrm{a})}{=}\boldsymbol{x}^{\mathrm{T}}\overline{A\boldsymbol{x}}\underset{(\mathrm{b})}{=}\boldsymbol{x}^{\mathrm{T}}A\overline{\boldsymbol{x}}\underset{(\mathrm{c})}{=}(\boldsymbol{x}^{\mathrm{T}}A\overline{\boldsymbol{x}})^{\mathrm{T}}\underset{(\mathrm{d})}{=}\overline{\boldsymbol{x}}^{\mathrm{T}}A^{\mathrm{T}}\boldsymbol{x}\underset{(\mathrm{e})}{=}q$$

24. 设 $n\times n$ 实矩阵 A 有性质 $A^{\mathrm{T}}=A$. 证明，若对 $\boldsymbol{x}\in\mathbb{C}^n$ 有 $A\boldsymbol{x}=\lambda\boldsymbol{x}$，则 λ 是实数，而 $\boldsymbol{x}$ 的实部是 A 的特征向量.（提示：计算 $\overline{\boldsymbol{x}}^{\mathrm{T}}A\boldsymbol{x}$ 和利用习题 23 的结果，并检查 $A\boldsymbol{x}$ 的实部和虚部.）

25. 设 A 是 $n\times n$ 实矩阵，$x\in\mathbb{C}^n$，证明 $\mathrm{Re}(A\boldsymbol{x})=A(\mathrm{Re}\,\boldsymbol{x})$ 和 $\mathrm{Im}(A\boldsymbol{x})=A(\mathrm{Im}\,\boldsymbol{x})$.

26. 设 A 是 2×2 实矩阵，有复特征值 $\lambda=a-b\mathrm{i}$ $(b\neq0)$ 和对应 $\mathbb{C}^2$ 中的复特征向量 $\boldsymbol{v}$，
   a. 证明 $A(\mathrm{Re}\,\boldsymbol{v})=a\,\mathrm{Re}\,\boldsymbol{v}+b\,\mathrm{Im}\,\boldsymbol{v}$ 和 $A(\mathrm{Im}\,\boldsymbol{v})=-b\,\mathrm{Re}\,\boldsymbol{v}+a\,\mathrm{Im}\,\boldsymbol{v}$.（提示:记 $v=\mathrm{Re}\,\boldsymbol{v}+\mathrm{i}\,\mathrm{Im}\,\boldsymbol{v}$，计算 $A\boldsymbol{v}$.）
   b. 假设 P 和 C 是定理 9 给出的矩阵，则 $AP=PC$.

[M]在习题 27 和习题 28 中，求所给矩阵 A 的分解式 $A=PCP^{-1}$，其中 C 是由形如例 6 矩阵的 2×2 子矩阵组成的分块对角矩阵.（对每一共轭对复特征值，利用属于 $\mathbb{C}^4$ 的特征向量的实部和虚部来产生 P 的两列.）

27. $A=\begin{bmatrix} 26 & 33 & 23 & 20 \\ -6 & -8 & -1 & -13 \\ -14 & -19 & -16 & 3 \\ -20 & -20 & -20 & -14 \end{bmatrix}$

28. $A=\begin{bmatrix} 7 & 11 & 20 & 17 \\ -20 & -40 & -86 & -74 \\ 0 & -5 & -10 & -10 \\ 10 & 28 & 60 & 53 \end{bmatrix}$

练习题答案

记住，要验证向量是否为特征向量是很容易的，并不需要检验特征方程. 计算
$$A\boldsymbol{x}=\begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} 1 \\ -\mathrm{i} \end{bmatrix}=\begin{bmatrix} a+b\mathrm{i} \\ b-a\mathrm{i} \end{bmatrix}=(a+b\mathrm{i})\begin{bmatrix} 1 \\ -\mathrm{i} \end{bmatrix}$$
因此，$\begin{bmatrix} 1 \\ -\mathrm{i} \end{bmatrix}$ 是对应 $\lambda=a+b\mathrm{i}$ 的特征向量，从本节讨论的内容知道，$\begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix}$ 一定是对应 $\overline{\lambda}=a-b\mathrm{i}$ 的特征向量.

5.6 离散动力系统

特征值和特征向量提供了线索，使我们理解由差分方程 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 描述的动力系统的长期行为或进化. 这种方程可用来建立人口动态变化的数学模型，如 1.10 节的人口变化模型，4.9 节的各种马尔可夫链及本章介绍性的斑点猫头鹰数学模型. 向量 $\boldsymbol{x}_k$ 给出系统随时间推移（记为 k）的相关信息. 例如，在斑点猫头鹰例子里，$\boldsymbol{x}_k$ 表示在时间 k 三个年龄段的猫头鹰的数目.

由于生态问题要比物理或工程上的问题容易描述和解释，本节的应用焦点放在生态问题上. 但很多的科学领域存在动力系统. 例如控制系统的标准大学课程对动力系统的某些方面进行了讨论.这些课程中的现代状态空间设计方法就主要依赖于矩阵代数[㊀]. 控制系统中的稳态响应在工程上等价于我们在这里所说的动力系统 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 的“长期行为”.

一直到例 6，我们假设 A 可对角化，有 n 个线性无关的特征向量 $v_1,\cdots,v_n$ 和对应的特征值 $\lambda_1,\cdots,\lambda_n$. 为方便起见，假设特征向量已按 $|\lambda_1|\geqslant|\lambda_2|\geqslant\cdots\geqslant|\lambda_n|$ 的顺序排列好. 因为 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$ 是 $\mathbb{R}^n$ 的基，故任一初始向量 x_0 可以唯一表示为

$$\boldsymbol{x}_0=c_1\boldsymbol{v}_1+\cdots+c_n\boldsymbol{v}_n \tag{1}$$

$\boldsymbol{x}_0$ 的这种特征向量分解确定了序列 $\{\boldsymbol{x}_k\}$ 所发生的情况. 下一步的计算将 5.2 节例 5 的简单情况一般化. 因为 $\boldsymbol{v}_i$ 是特征向量，所以

$$\begin{aligned}\boldsymbol{x}_1=A\boldsymbol{x}_0&=c_1A\boldsymbol{v}_1+\cdots+c_nA\boldsymbol{v}_n\\&=c_1\lambda_1\boldsymbol{v}_1+\cdots+c_n\lambda_n\boldsymbol{v}_n\end{aligned}$$

一般有

$$\boldsymbol{x}_k=c_1(\lambda_1)^k\boldsymbol{v}_1+\cdots+c_n(\lambda_n)^k\boldsymbol{v}_n\quad(k=0,1,2,\cdots) \tag{2}$$

下列例子说明当 $k\to\infty$ 时，（2）会出现什么结果.

捕食者-食饵系统

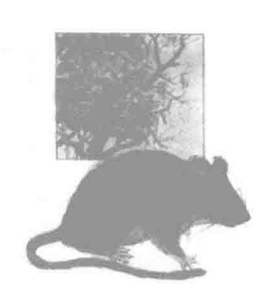
在加利福尼亚州的红木森林深处，作为老鼠的主要捕食者，斑点猫头鹰的食物有 80%是老鼠. 例1利用线性动力系统来建立猫头鹰和老鼠的自然系统模型.（实事求是地讲，这个模型在某些方面与现实不符，但它能够为环境科学家们所用的更复杂的非线性模型的研究提供一个起点.）

例 1 用 $\boldsymbol{x}_k=\begin{bmatrix}O_k\\R_k\end{bmatrix}$ 表示在时间 k（k 的单位是月）猫头鹰和老鼠的数量，O_k 是在研究区域猫头鹰的数量，R_k 是老鼠的数量（单位是千只）. 设

$$\begin{aligned}O_{k+1}&=(0.5)O_k+(0.4)R_k\\R_{k+1}&=-p\cdot O_k+(1.1)R_k\end{aligned} \tag{3}$$

㊀ 参见 G. F. Franklin, J. D. Powell, and A.Emami-Naeimi, *Feedback Control of Dynamic Systems*,5th ed.（Upper Saddle River,NJ: Prentice- Hall, 2006）. 这本大学教材对动力模型作了详细介绍（第 2 章）. 在第 7 章和第 8 章讨论状态空间设计.

其中 p 是被指定的正参数. 第 1 个方程中的 $(0.5)O_k$ 表示，如果没有老鼠为食物，每月仅有一半的猫头鹰存活下来，而第 2 个方程的 $(1.1)R_k$ 表明如果没有猫头鹰捕食老鼠，那么老鼠的数量每月增长 10%. 假如有足够多的老鼠，$(0.4)R_k$ 表示猫头鹰增长的数量，而负项 $-p\cdot O_k$ 表示由于猫头鹰的捕食所引起的老鼠的死亡数量.（事实上，一只猫头鹰每月平均吃掉 $1000p$ 只老鼠.）当 $p=0.104$ 时，预测该系统的发展趋势.



解 当 $p=0.104$ 时，算出方程组（3）的系数矩阵 A 的特征值是 $\lambda_1=1.02$ 和 $\lambda_2=0.58$. 对应的特征向量是

$$\boldsymbol{v}_1=\begin{bmatrix}10\\13\end{bmatrix},\ \boldsymbol{v}_2=\begin{bmatrix}5\\1\end{bmatrix}$$

初始向量 $\boldsymbol{x}_0$ 可表示为 $x_0=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2$，那么对 $k\geqslant 0$

$$\begin{aligned}\boldsymbol{x}_k&=c_1(1.02)^k\boldsymbol{v}_1+c_2(0.58)^k\boldsymbol{v}_2\\&=c_1(1.02)^k\begin{bmatrix}10\\13\end{bmatrix}+c_2(0.58)^k\begin{bmatrix}5\\1\end{bmatrix}\end{aligned}$$

当 $k\to\infty$ 时，$(0.58)^k$ 很快趋于零. 假设 $c_1>0$，那么对所有足够大的 k，$\boldsymbol{x}_k$ 近似等于 $c_1(1.02)^k\begin{bmatrix}10\\13\end{bmatrix}$，我们记为

$$\boldsymbol{x}_k\approx c_1(1.02)^k\begin{bmatrix}10\\13\end{bmatrix}\tag{4}$$

随着 k 的增大，（4）的近似程度会更好，故对足够大的 k

$$\boldsymbol{x}_{k+1}\approx c_1(1.02)^{k+1}\begin{bmatrix}10\\13\end{bmatrix}=(1.02)c_1(1.02)^k\begin{bmatrix}10\\13\end{bmatrix}\approx 1.02\boldsymbol{x}_k\tag{5}$$

近似式（5）表明最终 $\boldsymbol{x}_k$ 的 2 个分量（猫头鹰和老鼠的数量）每月以大约 1.02 的倍数增长，即月增长率为 2%. 由（4），$\boldsymbol{x}_k$ 就近似于（10,13）的倍数，因此，$\boldsymbol{x}_k$ 的 2 个分量之比率也近似于 10 与 13 的比率，也就是说，对应每 10 只猫头鹰，大致有 13 000 只老鼠. ■

例 1 说明了有关动力系统 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 的两个基本事实，若 A 是 $n\times n$ 矩阵，它的特征值满足 $|\lambda_1|\geqslant 1$ 和 $1>|\lambda_j|$，$j=2,\cdots,n$，$\boldsymbol{v}_1$ 是 λ_1 对应的特征向量，假如 $\boldsymbol{x}_0$ 由式（1）给出且 $c_1\neq 0$，那么对足够大的 k，

$$\boldsymbol{x}_{k+1}\approx\lambda_1\boldsymbol{x}_k\tag{6}$$

和

$$\boldsymbol{x}_k\approx c_1(\lambda_1)^k\boldsymbol{v}_1\tag{7}$$

式（6）和（7）的近似精度可根据需要通过取足够大的 k 来得到. 由式（6），$\boldsymbol{x}_k$ 每时段最终以近似 λ_1 的倍数增长，因此，λ_1 确定了系统的最终增长率. 同样由式（7），对足够大的 k，$\boldsymbol{x}_k$ 的 2 个分量之比近似等于 $\boldsymbol{v}_1$ 对应分量之比. 5.2 节的例 5 是 $\lambda_1=1$ 的实例.

解的几何意义

当 A 为 2×2 矩阵时，可以通过系统发展趋势的几何描述来补充解释代数计算. 我们可以把

方程 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 看作是 $\mathbb{R}^2$ 中的初始点 x_0 被映射 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 重复变换的描述，由 $\boldsymbol{x}_0,\boldsymbol{x}_1,\cdots$ 组成的图形称为是动力系统的**轨迹**.

例 2 当 $A=\begin{bmatrix}0.80 & 0\\ 0 & 0.64\end{bmatrix}$ 时，画出动力系统 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 的若干条轨迹.

解 A 的特征值是 0.8 和 0.64，对应的特征向量是 $\boldsymbol{v}_1=\begin{bmatrix}1\\0\end{bmatrix}$ 和 $\boldsymbol{v}_2=\begin{bmatrix}0\\1\end{bmatrix}$. 假如 $\boldsymbol{x}_0=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2$，那么

$$\boldsymbol{x}_k=c_1(0.8)^k\begin{bmatrix}1\\0\end{bmatrix}+c_2(0.64)^k\begin{bmatrix}0\\1\end{bmatrix}$$

当 $k\to\infty$ 时，$(0.8)^k$ 和 $(0.64)^k$ 都趋于零，当然 $\boldsymbol{x}_k$ 也趋于零. 但 $\boldsymbol{x}_k$ 趋于零的方式是有趣的. 图 5-16 显示了几条轨迹的开头几项，这些轨迹的起点在四个角点的坐标为 $(\pm3,\pm3)$ 的矩形的边界上. 为使轨迹容易看清，用细线把轨迹上的点连接起来. ■

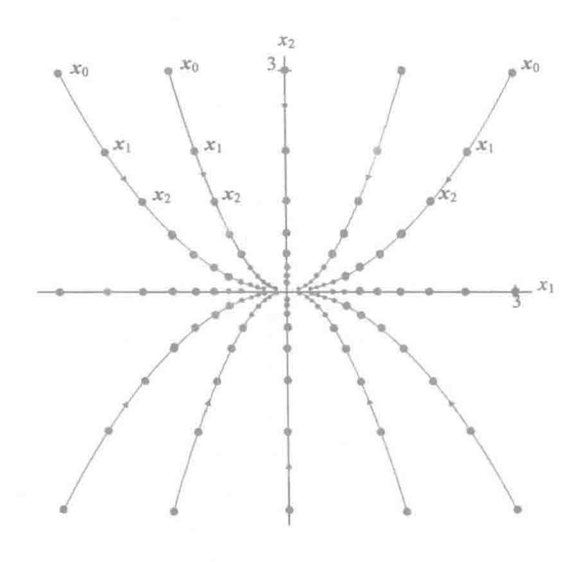


图 5-16 原点是吸引子

在例 2，因为所有的轨迹都趋于原点，所以原点被称作动力系统的**吸引子**. 当两个特征值的绝对值都小于 1 的时候出现这种情况. 过原点和最小绝对值的特征值的特征向量 v_2 的直线的方向是最大吸引方向.

在下一个例子，A 的两个特征值的绝对值都大于 1，此时的原点称为动力系统的**排斥子**. 除了（常数）零解，$\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 的所有解是无界的，离原点而去.[⊖]

⊖ 在线性动力系统中，只有原点才可能是吸引子或排斥子，但在映射 $\boldsymbol{x}_k\mapsto\boldsymbol{x}_{k+1}$ 为非线性的更一般的动力系统中，可能存在多个吸引子和排斥子. 在这样的系统中，吸引子和排斥子用某个特殊矩阵（有变量元素）的特征值来定义，这个矩阵称为系统的雅可比矩阵.

例 3 画出方程 $x_{k+1}=Ax_k$ 的若干条典型轨迹，这里

$$A=\begin{bmatrix}1.44 & 0\\ 0 & 1.2\end{bmatrix}$$

解 A 的特征值是 1.44 和 1.2，若 $x_0=\begin{bmatrix}c_1\\ c_2\end{bmatrix}$，则

$$x_k=c_1(1.44)^k\begin{bmatrix}1\\ 0\end{bmatrix}+c_2(1.2)^k\begin{bmatrix}0\\ 1\end{bmatrix}$$

两项的值随 k 增大而增大，但第 1 项增大的快一些. 因此，过原点和较大特征值的特征向量的直线方向是最大排斥方向. 图 5-17 显示的是起点接近原点的几条轨迹.

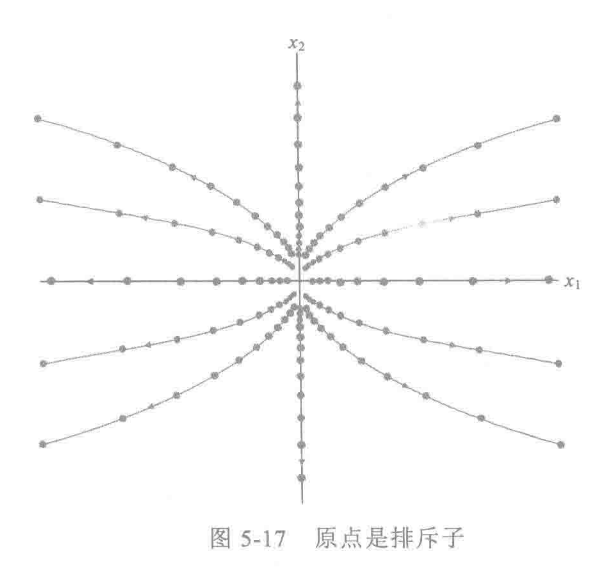


图 5-17 原点是排斥子 ■

在下一个例子，原点称为**鞍点**，因为原点在某些方向吸引解，而在其他方向又排斥解. 当一个特征值的绝对值大于 1 而另一个特征值的绝对值小于 1 的时候出现这种情况. 最大的吸引方向是由绝对值较小的特征值的特征向量决定的. 最大的排斥方向是由绝对值较大的特征值的特征向量决定的.

例 4 画出方程 $y_{k+1}=Dy_k$ 的若干典型解的轨迹，其中

$$D=\begin{bmatrix}2.0 & 0\\ 0 & 0.5\end{bmatrix}$$

（我们这里用 D 和 y 代替 A 和 x 是因为后面要用到该例.）

解 D 的特征值是 2 和 0.5. 若 $y_0=\begin{bmatrix}c_1\\ c_2\end{bmatrix}$，那么

$$y_k=c_1 2^k\begin{bmatrix}1\\ 0\end{bmatrix}+c_2(0.5)^k\begin{bmatrix}0\\ 1\end{bmatrix} \tag{8}$$

假如 y_0 在 x_2 轴，那么 $c_1=0$，因此当 $k\to\infty$ 时，$y_k\to\mathbf{0}$. 但当 y_0 不在 x_2 轴时，计算 y_k 的和式中的第 1 项变得任意大，因此 $\{y_k\}$ 是无界的. 图 5-18 显示起点靠近或在 x_2 轴上的 10 条轨迹.

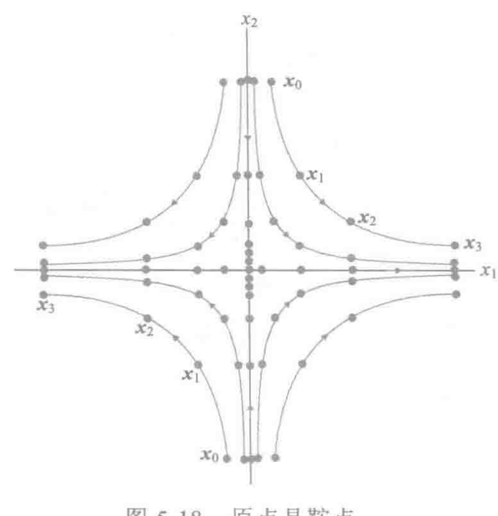


图 5-18 原点是鞍点 ■

变量代换

前面 3 个例子讨论的矩阵是对角矩阵，为处理非对角矩阵，我们先暂时回到 A 为 $n\times n$ 矩阵的情形，设 A 的特征向量 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$ 是 $\mathbb{R}^n$ 的基.令 $P=[\boldsymbol{v}_1\cdots\boldsymbol{v}_n]$，$D$ 是对角线上元素为对应特征值的对角矩阵.给出序列 $\{\boldsymbol{x}_k\}$ 满足 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$，由

$$\boldsymbol{y}_k=P^{-1}\boldsymbol{x}_k \text{ 或 } \boldsymbol{x}_k=P\boldsymbol{y}_k$$

定义一个新的序列 $\{\boldsymbol{y}_k\}$，把这些关系代入方程 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$，并利用 $A=PDP^{-1}$，我们求得

$$P\boldsymbol{y}_{k+1}=AP\boldsymbol{y}_k=(PDP^{-1})P\boldsymbol{y}_k=PD\boldsymbol{y}_k$$

两边乘 P^{-1}，得

$$\boldsymbol{y}_{k+1}=D\boldsymbol{y}_k$$

假如我们记 $\boldsymbol{y}_k$ 为 $\boldsymbol{y}(k)$，用 $y_1(k),\cdots,y_n(k)$ 表示 $\boldsymbol{y}(k)$ 的分量，那么

$$\begin{bmatrix} y_1(k+1) \\ y_2(k+1) \\ \vdots \\ y_n(k+1) \end{bmatrix}=\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}\begin{bmatrix} y_1(k) \\ y_2(k) \\ \vdots \\ y_n(k) \end{bmatrix}$$

从 $\boldsymbol{x}_k$ 到 $\boldsymbol{y}_k$ 的变量代换解耦了差分方程系统. 例如，$y_1(k)$ 的变化不受 $y_2(k),\cdots,y_n(k)$ 的影响，因为对每一个 $k, y_1(k+1)=\lambda_1 y_1(k)$.

等式 $\boldsymbol{x}_k=P\boldsymbol{y}_k$ 表明 $\boldsymbol{y}_k$ 是 $\boldsymbol{x}_k$ 在向量基 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$ 下的坐标向量. 这样我们就可以通过在新的向量坐标系中进行计算来解耦系统 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$，当 $n=2$ 时，相当于在使用坐标轴在两个特征向量方向上的方格纸.

例 5 证明：原点是方程 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 解的鞍点，其中

$$A=\begin{bmatrix} 1.25 & -0.75 \\ -0.75 & 1.25 \end{bmatrix}$$

并求最大的吸引方向和排斥方向.

解 用通常方法，可以求得 A 有特征值 2 和 0.5，对应的特征向量分别是

$$\boldsymbol{v}_1=\begin{bmatrix}1\\-1\end{bmatrix} \text{和} \ \boldsymbol{v}_2=\begin{bmatrix}1\\1\end{bmatrix}$$

由于 $|2|>1$ 和 $|0.5|<1$，因此原点是动力系统的鞍点，假如 $\boldsymbol{x}_0=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2$，那么

$$\boldsymbol{x}_k=c_1 2^k\boldsymbol{v}_1+c_2(0.5)^k\boldsymbol{v}_2 \tag{9}$$

这个等式看起来像例 4 的式（8），只是用 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 代替了标准基.

在方格纸上，过 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 画坐标轴. 看图 5-19，沿这些轴的移动相当于例 3 的沿标准轴的移动. 在图 5-19，最大的排斥方向是在过 $\boldsymbol{v}_1$ 的直线上，因为 $\boldsymbol{v}_1$ 对应的特征值大于 1. 若 $\boldsymbol{x}_0$ 在这条直线上，则（9）中的 $c_2=0$，因此 $\boldsymbol{x}_k$ 快速远离原点，最大的吸引方向由特征向量 $\boldsymbol{v}_2$ 确定，$\boldsymbol{v}_2$ 对应的特征值小于 1. 图 5-19 显示了一些轨迹. 若按特征向量轴来看这些图，图的形状与例 3 的一样.

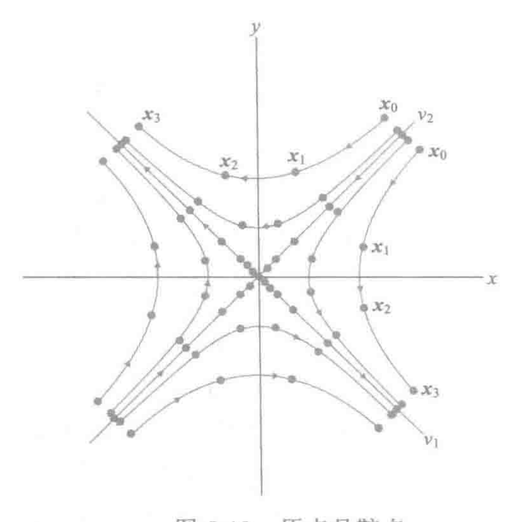


图 5-19 原点是鞍点 ■

复特征值

若 A 为有复特征值的 2×2 矩阵，则 A 不可对角化（当 A 作用在 $\mathbb{R}^n$ 时），但动力系统 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 还是容易描述. 5.5 节的例 3 给出了这样的例子，此时，特征值的长度为 1. 点 $\boldsymbol{x}_0$ 的迭代绕原点沿椭圆型轨道作螺旋运动.

假如 A 有两个长度都大于 1 的复特征值，那么原点是排斥子，$\boldsymbol{x}_0$ 的迭代绕原点向外作螺旋线旋转. 假如复特征值的长度都小于 1，原点是吸引子，$\boldsymbol{x}_0$ 的迭代绕原点向内作螺旋线旋转. 见下例.

例 6 可以验证矩阵

$$A=\begin{bmatrix}0.8 & 0.5\\-0.1 & 1.0\end{bmatrix}$$

有特征值 $0.9 \pm 0.2\mathrm{i}$，对应的特征向量是 $\begin{bmatrix} 1 \mp 2\mathrm{i} \\ 1 \end{bmatrix}$，图 5-20 显示了初始向量为 $\begin{bmatrix} 0 \\ 2.5 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \end{bmatrix}$ 和 $\begin{bmatrix} 0 \\ -2.5 \end{bmatrix}$ 时动力系统 $\boldsymbol{x}_{k+1} = A\boldsymbol{x}_k$ 的 3 条轨迹.

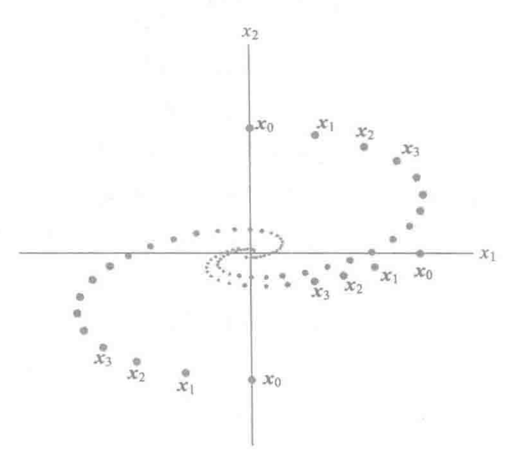


图 5-20　与复特征值相关的旋转

斑点猫头鹰的生存

回顾本章介绍性实例，我们使用动力系统 $\boldsymbol{x}_{k+1} = A\boldsymbol{x}_k$ 为 California Willow Creek 的猫头鹰建立种群模型，在该模型中，$\boldsymbol{x}_k = (j_k, s_k, a_k)$ 的分量分别表示在时间 k 幼年、半成年和成年雌性猫头型的数量，A 为阶段矩阵

$$A = \begin{bmatrix} 0 & 0 & 0.33 \\ 0.18 & 0 & 0 \\ 0 & 0.71 & 0.94 \end{bmatrix} \tag{10}$$

由 MATLAB 算出 A 的特征值大约是 $\lambda_1 = 0.98, \lambda_2 = -0.02 + 0.21\mathrm{i}$ 和 $\lambda_3 = -0.02 - 0.21\mathrm{i}$，因为 $|\lambda_2|^2 = |\lambda_3|^2 = (-0.02)^2 + (0.21)^2 = 0.044\,5$，故三个特征值的长度都小于 1.

现在，让 A 作用在复空间 $\mathbb{C}^3$. 因为 A 有 3 个相异的特征值，故对应的 3 个特征向量是线性无关的，形成 $\mathbb{C}^3$ 的一个基. 记这些特征向量为 $\boldsymbol{v}_1, \boldsymbol{v}_2$ 和 $\boldsymbol{v}_3$. 那么 $\boldsymbol{x}_{k+1} = A\boldsymbol{x}_k$ 的通解（用 $\mathbb{C}^3$ 的向量表示）的形式是

$$x_k = c_1(\lambda_1)^k \boldsymbol{v}_1 + c_2(\lambda_2)^k \boldsymbol{v}_2 + c_3(\lambda_3)^k \boldsymbol{v}_3 \tag{11}$$

若初始向量 $\boldsymbol{x}_0$ 是实向量，由于 A 是实矩阵，故 $\boldsymbol{x}_1 = A\boldsymbol{x}_0$ 也是实向量. 同样，方程 $\boldsymbol{x}_{k+1} = A\boldsymbol{x}_k$ 表明式（11）左边的 $\boldsymbol{x}_k$ 也是实向量，尽管它表示为复向量的和，但由于所有特征值都小于 1，所以式（11）右边的每一项都趋于零向量. 因此，实序列 $\{\boldsymbol{x}_k\}$ 也趋于零向量. 很不幸，该模型预测斑点猫头鹰会最终全部灭亡.

猫头鹰还有希望吗？回顾本章介绍性实例，在（10）的矩阵 A 中的元素 18%源于如下事实：尽管有 60%的幼年猫头鹰能够活下来，离巢去寻找新的栖息地，但在这 60%的幼年猫头鹰中，仅有 30%的猫头鹰能活下来找到新的栖息地. 森林中裸露地域的数量使得搜寻工作变得更困难和更危险，严重影响了寻找栖息地过程中的猫头鹰的存活率.

有相当数量的猫头鹰生活在没有或很少裸露地域的地方，这使得有更大百分比的幼年猫头鹰能够找到新的栖息地存活下来. 当然，猫头鹰的问题比我们这里描述的还要复杂，但最后一个例子将会给你一个满意的结果.

例 7 设幼年猫头鹰的搜寻存活率是 50%，因此（10）式中的矩阵 A 的（2,1）元素是 0.3 而不是 0.18，用这样的阶段矩阵模型预测猫头鹰数量的发展趋势.

解 现在 A 的特征值大约是 $\lambda_1 = 1.01, \lambda_2 = -0.03+0.26\mathrm{i}$ 和 $\lambda_3 = -0.03-0.26\mathrm{i}$，对应 λ_1 的特征向量为 $\boldsymbol{v}_1 = (10,3,31)$，并设 $\boldsymbol{v}_2$ 和 $\boldsymbol{v}_3$ 是对应 λ_2 和 λ_3 的（复）特征向量. 此时，等式（11）变为

$$\boldsymbol{x}_k = c_1(1.01)^k\boldsymbol{v}_1 + c_2(-0.03+0.26\mathrm{i})^k\boldsymbol{v}_2 + c_3(-0.03-0.26\mathrm{i})^k\boldsymbol{v}_3$$

当 $k\to\infty$ 时，向量 $\boldsymbol{v}_2,\boldsymbol{v}_3$ 趋于零，因此 $\boldsymbol{x}_k$ 越来越接近（实）向量 $c_1(1.01)^k\boldsymbol{v}_1$. 例 1 中的近似值（6）和（7）在这里仍适用，因此，猫头鹰数量的长期增长率是 1.01，即猫头鹰的数量会缓慢增长. 特征向量 $\boldsymbol{v}_1$ 描述了猫头鹰在 3 个年龄段数量的最终分布，每 31 只成年猫头鹰，对应大约有 10 只幼年猫头鹰和 3 只半成年猫头鹰. ■

进一步阅读

Franklin, G. F., J. D. Powell, and M. L. Workman. *Digital Control of Dynamic Systems*, 3rd ed. Reading, MA: Addison-Wesley, 1998.

Sandefur, James T. *Discrete Dynamical Systems—Theory and Applications*. Oxford: Oxford University Press, 1990.

Tuchinsky, Philip. *Management of a Buffalo Herd*, UMAP Module 207. Lexington, MA: COMAP, 1980.

练习题

1. 下面矩阵 A 的特征值是 1，2/3 和 1/3，相应的特征向量分别为 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$：

$$A=\frac{1}{9}\begin{bmatrix}7&-2&0\\-2&6&2\\0&2&5\end{bmatrix}, \boldsymbol{v}_1=\begin{bmatrix}-2\\2\\1\end{bmatrix}, \boldsymbol{v}_2=\begin{bmatrix}2\\1\\2\end{bmatrix}, \boldsymbol{v}_3=\begin{bmatrix}1\\2\\-2\end{bmatrix}$$

若 $\boldsymbol{x}_0=\begin{bmatrix}1\\11\\-2\end{bmatrix}$，求方程 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 的通解.

2. 在练习题 1 中，当 $k\to\infty$ 时，序列 $\{\boldsymbol{x}_k\}$ 是什么？

习题 5.6

1. 设 2×2 矩阵 A 的特征值是 3 和 1/3，对应的特征向量为 $\boldsymbol{v}_1=\begin{bmatrix}1\\1\end{bmatrix}$ 和 $\boldsymbol{v}_2=\begin{bmatrix}-1\\1\end{bmatrix}$. $\{\boldsymbol{x}_k\}$ 是差分方程 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 的解，$\boldsymbol{x}_0=\begin{bmatrix}9\\1\end{bmatrix}$.

   a. 计算 $\boldsymbol{x}_1=A\boldsymbol{x}_0$.（提示：你不需要知道 A 本身.）

   b. 求 $\boldsymbol{x}_k$ 包含 k 和特征向量 $\boldsymbol{v}_1$ 及 $\boldsymbol{v}_2$ 的公式.

2. 假设 3×3 矩阵 A 的特征值是 3，4/5，3/5，对应的特征向量为 $\begin{bmatrix}1\\0\\-3\end{bmatrix},\begin{bmatrix}2\\1\\-5\end{bmatrix},\begin{bmatrix}-3\\-3\\7\end{bmatrix}$. 设 $\boldsymbol{x}_0=\begin{bmatrix}-2\\-5\\3\end{bmatrix}$，对给定的 $\boldsymbol{x}_0$ 求方程 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 的解，并描述当 $k\to\infty$ 时有何结果.

在习题 3～6 中，假设任意初始向量 $\boldsymbol{x}_0$ 有一个满

足本节式（1）中系数 c_1 为正的特征向量分解. ⊖

3. 在例 1 的动力系统中，若方程组（3）的捕食参数 $p=0.2$，预测动力系统的发展趋势（给出 $\boldsymbol{x}_k$ 的公式）. 猫头鹰的数量是增长还是下降呢？老鼠的情况又怎样？
4. 在例 1 的动力系统中，若捕食参数 p=0.125，预测动力系统的发展趋势（给出 $\boldsymbol{x}_k$ 的公式）. 随着时间的推移，猫头鹰和老鼠的数量会发生什么变化？这种系统趋势有时称为不稳定平衡. 如果这个模型的某些方面（如出生率或捕食率）稍微发生改变，这个系统会怎样？
5. 在古老的 Douglas 冷杉森林中，斑点猫头鹰主要以鼯鼠为食. 设这两个种群的捕食者-食饵矩阵 $A=\begin{bmatrix}0.4 & 0.3\\ -p & 1.2\end{bmatrix}$，证明：若捕食参数 p 为 0.325，则两个种群的数量都是增长的. 预测长期增长率及猫头鹰与鼯鼠的最终比率.
6. 若习题 5 的捕食参数为 0.5，证明猫头鹰和鼯鼠最终都会灭亡. p 取何值时，两者的数量保持稳定？此时，两者的数量关系是什么？
7. 设 A 具有习题 1 描述的性质.
   a. 原点是动力系统 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 的吸引子还是排斥子还是鞍点？
   b. 求该动力系统的最大吸引方向或排斥方向.
   c. 作该系统的几何描述，显示最大吸引方向或排斥方向，包括若干典型轨迹的草图（不用计算具体的点）.
8. 若 A 具有习题 2 描述的性质，原点是该动力系统 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 的吸引子、排斥子还是鞍点？求最大的吸引方向或排斥方向.

在习题 9~14 中，把原点归类为动力系统 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 的吸引子、排斥子或鞍点，并求最大的吸引方向或排斥方向.

9. $A=\begin{bmatrix}1.7 & -0.3\\ -1.2 & 0.8\end{bmatrix}$
10. $A=\begin{bmatrix}0.3 & 0.4\\ -0.3 & 1.1\end{bmatrix}$
11. $A=\begin{bmatrix}0.4 & 0.5\\ -0.4 & 1.3\end{bmatrix}$
12. $A=\begin{bmatrix}0.5 & 0.6\\ -0.3 & 1.4\end{bmatrix}$
13. $A=\begin{bmatrix}0.8 & 0.3\\ -0.4 & 1.5\end{bmatrix}$
14. $A=\begin{bmatrix}1.7 & 0.6\\ -0.4 & 0.7\end{bmatrix}$
15. 设 $A=\begin{bmatrix}0.4 & 0 & 0.2\\ 0.3 & 0.8 & 0.3\\ 0.3 & 0.2 & 0.5\end{bmatrix}$，向量 $\boldsymbol{v}_1=\begin{bmatrix}0.1\\ 0.6\\ 0.3\end{bmatrix}$ 是 A 的特征向量，A 的两个特征值是 0.5 和 0.2，求动力系统 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 满足 $\boldsymbol{x}_0=(0,0.3,0.7)$ 的解，当 $k\to\infty$ 时，$\boldsymbol{x}_k$ 会如何？
16. [M] A 为 4.9 节习题 16 中 Hertz 租车模型的随机矩阵，求动力系统 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 的通解.
17. 为某动物种类建立阶段矩阵模型. 该动物的生命周期分 2 个阶段，幼年期（1 岁以前）和成年期. 假设每只成年雌性一年平均生下 1.6 只幼年雌性. 每年，有 30%的幼年存活下来进入成年和 80%的成年仍然存活. 对 $k\geqslant 0$，设 $\boldsymbol{x}_k=(j_k,a_k)$，其中 $\boldsymbol{x}_k$ 的分量表示在 k 年幼年和成年的数量.
    a. 构造阶段矩阵 A，使得对 $k\geqslant 0$ 时，有 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$.
    b. 证明动物的数量的是增长的，并计算最终增长率和幼年与成年的最终比率.
    c. [M]假设种群最初有 15 只幼年和 10 只成年. 画 4 个图显示在未来 8 年种群数量的变化情况：(a) 幼年的数量，(b) 成年数量，(c) 总数量，(d) 幼年与成年的比率（每年）. 什么时候（d）中的比率会达到稳定？要求列出产生（c）图和（d）图的程序或命令.
18. 可以用类似斑点猫头鹰的阶段矩阵为美国野牛群建立模型. 雌性野牛被分为：小牛（1 岁以前），半成年野牛（1~2 岁）和成年野牛. 假设每 100 头成年雌性野牛每年平均生下 42 头雌性

⊖ 例 1 中模型的一种极限是总是存在初始向量 $\boldsymbol{x}_0$，其分量为正，使得系数 c_1 为负. 近似式（7）仍然成立，但 $\boldsymbol{x}_k$ 的分量为负.

小牛（只有成年雌性野牛能够产崽）．每年大约有 60%的小牛，75%半成年野牛和 95%成年野牛存活．对 $k \geqslant 0$，令 $\boldsymbol{x}_k = (c_k, y_k, a_k)$，$\boldsymbol{x}_k$ 的分量表示在 k 年三个年龄段雌性野牛的数量．

a. 构造野牛群的阶段矩阵 A，使得对 $k \geqslant 0, \boldsymbol{x}_{k+1} = A\boldsymbol{x}_k$．

b. [M]证明野牛群的数量是增长的，预测若干年后的增长率和每 100 头成年野牛对应的小牛和半成年野牛的数量．

练习题答案

1\. 第 1 步将 $\boldsymbol{x}_0$ 表示为 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ 的线性组合．把 $[\boldsymbol{v}_1 \ \ \boldsymbol{v}_2 \ \ \boldsymbol{v}_3 \ \ \boldsymbol{x}_0]$ 行化简，求出系数 $c_1 = 2, c_2 = 1, c_3 = 3$，因此

$$\boldsymbol{x}_0 = 2\boldsymbol{v}_1 + 1\boldsymbol{v}_2 + 3\boldsymbol{v}_3$$

由于特征值是 $1, \frac{2}{3}$ 和 $\frac{1}{3}$，故通解是

$$\boldsymbol{x}_k = 2 \cdot 1^k \boldsymbol{v}_1 + 1 \cdot \left(\frac{2}{3}\right)^k \boldsymbol{v}_2 + 3 \cdot \left(\frac{1}{3}\right)^k \boldsymbol{v}_3 = 2\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} + \left(\frac{2}{3}\right)^k \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + 3 \cdot \left(\frac{1}{3}\right)^k \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} \tag{12}$$

2\. 当 $k \to \infty$ 时，（12）中的第 2 和第 3 项趋于零向量，故

$$\boldsymbol{x}_k = 2\boldsymbol{v}_1 + \left(\frac{2}{3}\right)^k \boldsymbol{v}_2 + 3\left(\frac{1}{3}\right)^k \boldsymbol{v}_3 \to 2\boldsymbol{v}_1 = \begin{bmatrix} -4 \\ 4 \\ 2 \end{bmatrix}$$

5.7 微分方程中的应用

本节讲述在 5.6 节研究的差分方程的连续型类推．在很多的应用问题中，有些量随时间连续变化，它们与下面的微分方程组有关：

$$\begin{aligned} x_1' &= a_{11}x_1 + \cdots + a_{1n}x_n \\ x_2' &= a_{21}x_1 + \cdots + a_{2n}x_n \\ &\vdots \\ x_n' &= a_{n1}x_1 + \cdots + a_{nn}x_n \end{aligned}$$

这里 $x_1, \cdots, x_n$ 是 t 的可导函数，导数分别是 $x_1', \cdots, x_n'$，a_{ij} 是常数．该方程组最主要的特征是线性性质．为了便于理解，我们把方程组写成矩阵微分方程

$$\boldsymbol{x}'(t) = A\boldsymbol{x}(t) \tag{1}$$

其中

$$\boldsymbol{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \ \boldsymbol{x}'(t) = \begin{bmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{bmatrix}, \ A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

方程（1）的解是向量值函数，该函数定义在某实数区间，比如 $t \geqslant 0$，且满足方程（1）．

由于函数求导以及向量与矩阵相乘都是线性变换，故方程（1）是线性的．因此，若 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是 $\boldsymbol{x}' = A\boldsymbol{x}$ 的解，则 $c\boldsymbol{u} + d\boldsymbol{v}$ 同样也是 $\boldsymbol{x}' = A\boldsymbol{x}$ 的解，因为

$$(c\boldsymbol{u} + d\boldsymbol{v})' = c\boldsymbol{u}' + d\boldsymbol{v}' = cA\boldsymbol{u} + dA\boldsymbol{v} = A(c\boldsymbol{u} + d\boldsymbol{v})$$

（工程师们将这个性质称为解的**叠加**）同样，恒等于零的函数也是方程（1）的（平凡）解．用第 4 章的术语，方程（1）的所有解的集合是值属于 $\mathbb{R}^n$ 的所有连续函数组成的集合的子空间．

有关微分方程的标准教材证明了方程（1）一定存在**基础解系**. 假如 A 是 $n\times n$ 矩阵，那么在基本解集中存在 n 个线性无关的函数. 使得方程（1）的每一个解可以唯一表示为这 n 个函数的线性组合. 即基础解系是方程（1）的所有解的集合的基. 若给定向量 $\boldsymbol{x}_0$，那么**初值问题**就是构造一个（唯一）函数 $\boldsymbol{x}$，满足 $\boldsymbol{x}'=A\boldsymbol{x}$ 和 $\boldsymbol{x}(0)=\boldsymbol{x}_0$.

当 A 是对角矩阵时，可以用初等微积分求出（1）的解. 例如考虑

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix}=\begin{bmatrix} 3 & 0 \\ 0 & -5 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{2}$$

即有

$$\begin{aligned} x_1'(t) &= 3x_1(t) \\ x_2'(t) &= -5x_2(t) \end{aligned} \tag{3}$$

因为每个函数的导数仅依赖于自身，而不是 $x_1(t)$ 和 $x_2(t)$ 的组合或“结合”，此时称方程组（2）是解耦的. 由微积分，（3）的解是 $x_1(t)=c_1\mathrm{e}^{3t}$ 和 $x_2(t)=c_2\mathrm{e}^{-5t}$，$c_1$ 和 c_2 为任意常数.（2）的每一个解都可以写成下列形式：

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}=\begin{bmatrix} c_1\mathrm{e}^{3t} \\ c_2\mathrm{e}^{-5t} \end{bmatrix}=c_1\begin{bmatrix} 1 \\ 0 \end{bmatrix}\mathrm{e}^{3t}+c_2\begin{bmatrix} 0 \\ 1 \end{bmatrix}\mathrm{e}^{-5t}$$

这个例子提示我们，对于一般的方程 $\boldsymbol{x}'=A\boldsymbol{x}$，它的解可能是形如

$$\boldsymbol{x}(t)=\boldsymbol{v}\mathrm{e}^{\lambda t} \tag{4}$$

的函数的线性组合，其中 λ 为数，$\boldsymbol{v}$ 为非零向量.（若 $\boldsymbol{v}=\boldsymbol{0}$，函数 $\boldsymbol{x}(t)$ 恒为零，且满足 $\boldsymbol{x}'=A\boldsymbol{x}$.）注意到

$$\boldsymbol{x}'(t)=\lambda\boldsymbol{v}\mathrm{e}^{\lambda t} \qquad \text{对 } \boldsymbol{x}(t) \text{ 求导，其中 } \boldsymbol{v} \text{ 是常向量}$$

$$A\boldsymbol{x}(t)=A\boldsymbol{v}\mathrm{e}^{\lambda t} \qquad \text{式（4）两边同乘 } A$$

因为 $\mathrm{e}^{\lambda t}$ 不可能为零，故 $\boldsymbol{x}'(t)$ 等于 $A\boldsymbol{x}(t)$ 当且仅当 $\lambda\boldsymbol{v}=A\boldsymbol{v}$，即当且仅当 λ 是 A 的特征值，而 $\boldsymbol{v}$ 是对应的特征向量. 因此，每一对特征值-特征向量提供了 $\boldsymbol{x}'=A\boldsymbol{x}$ 的一个解（4），这种解有时被称为微分方程的**特征函数**. 特征函数为求解微分方程提供了方法.

例 1 图 5-21 显示的电路可以用微分方程描述

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix}=\begin{bmatrix} -(1/R_1+1/R_2)/C_1 & 1/(R_2C_1) \\ 1/(R_2C_2) & -1/(R_2C_2) \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

其中 $\boldsymbol{x}_1(t)$ 和 $\boldsymbol{x}_2(t)$ 是在时间 t 两个电容器的电压，设电阻 R_1 为 1 欧姆，R_2 为 2 欧姆，电容器 C_1 为 1 法拉，C_2 为 0.5 法拉，并假设电容器 C_1 的初始电压为 5 伏，C_2 为 4 伏. 求描述电压随时间变化的公式 $\boldsymbol{x}_1(t)$ 和 $\boldsymbol{x}_2(t)$.

图 5-21

解 由给出的数据，令 $A=\begin{bmatrix} -1.5 & 0.5 \\ 1 & -1 \end{bmatrix}$，$\boldsymbol{x}(t)=\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$，$\boldsymbol{x}_{(0)}=\begin{bmatrix} 5 \\ 4 \end{bmatrix}$，我们可以求得 A 的特征值是 $\lambda_1=-0.5$ 和 $\lambda_2=-2$，对应的特征向量是

$$\boldsymbol{v}_1=\begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{和} \quad \boldsymbol{v}_2=\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

特征函数 $\boldsymbol{x}_1(t)=\boldsymbol{v}_1\mathrm{e}^{\lambda_1 t}$ 和 $\boldsymbol{x}_2(t)=\boldsymbol{v}_2\mathrm{e}^{\lambda_2 t}$ 都满足 $\boldsymbol{x}'=A\boldsymbol{x}$，$\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 的任意线性组合亦同样满足 $\boldsymbol{x}'=A\boldsymbol{x}$. 令

$$x(t)=c_1v_1e^{\lambda_1 t}+c_2v_2e^{\lambda_2 t}=c_1\begin{bmatrix}1\\2\end{bmatrix}e^{-0.5t}+c_2\begin{bmatrix}-1\\1\end{bmatrix}e^{-2t}$$

记 $x(0)=c_1v_1+c_2v_2$，显然 v_1 和 v_2 是线性无关的，故 v_1 和 v_2 可生成 $\mathbb{R}^2$，令 $x(0)=x_0$，可求出 c_1 和 c_2. 事实上，由方程

$$\underset{\substack{\uparrow\\ v_1}}{c_1\begin{bmatrix}1\\2\end{bmatrix}}+\underset{\substack{\uparrow\\ v_2}}{c_2\begin{bmatrix}-1\\1\end{bmatrix}}=\underset{\substack{\uparrow\\ x_0}}{\begin{bmatrix}5\\4\end{bmatrix}}$$

容易解出 $c_1=3$ 和 $c_2=-2$，因此，微分方程 $x'=Ax$ 的解是

$$x(t)=3\begin{bmatrix}1\\2\end{bmatrix}e^{-0.5t}-2\begin{bmatrix}-1\\1\end{bmatrix}e^{-2t}$$

或

$$\begin{bmatrix}x_1(t)\\x_2(t)\end{bmatrix}=\begin{bmatrix}3e^{-0.5t}+2e^{-2t}\\6e^{-0.5t}-2e^{-2t}\end{bmatrix}$$

图 5-22 显示了 $x(t)$ 在 $t\geqslant 0$ 的图像或轨迹，一起显示的还有其他初始点的轨迹. 两个特征函数 x_1 和 x_2 的轨迹包含在 A 的特征空间里.

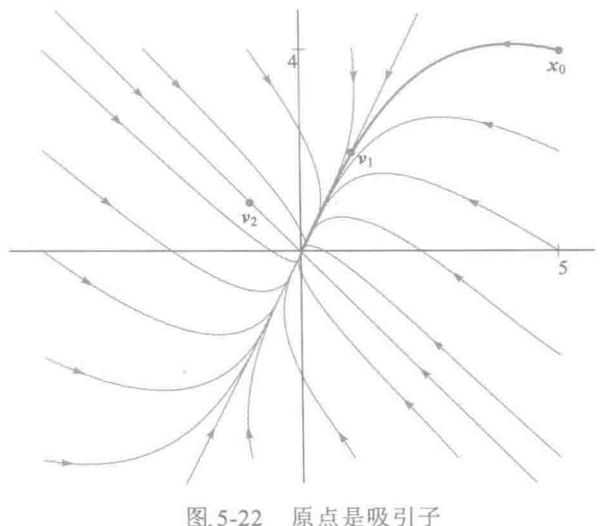


图 5-22　原点是吸引子 ■

当 $t\to\infty$ 时，函数 x_1 和 x_2 都衰减为零，但 x_2 的值要衰减得更快一些，因为它的指数要小一些. 对应的特征向量 v_2 的分量表明，若两个初始电压大小相等但符号相反，两个电容器的电压将会很快衰退为零.

在图 5-22 中，因为所有轨迹都趋近于原点，所以把原点称为动力系统的**吸引子**或**汇**. 最大的吸引方向是在较小的特征值 $\lambda=-2$ 对应的特征函数 x_2 的轨迹上（沿着过原点和 v_2 的直线）. 起点不在此直线的轨迹渐渐逼近过原点和 v_1 的直线，因为它们在 v_2 方向的分量衰减得很快.

如果例 1 的特征值是正数，相应的轨迹形状相同，但轨迹背离原点. 此时，称原点为动力

系统的**排斥子**或**源**，最大的排斥方向是在包含较大特征值对应的特征函数的轨迹的直线上.

例 2 假设粒子在平面力场中运动，它的位置向量 $\boldsymbol{x}$ 满足 $\boldsymbol{x}'=A\boldsymbol{x}$ 和 $\boldsymbol{x}(0)=\boldsymbol{x}_0$，其中

$$A=\begin{bmatrix}4 & -5\\ -2 & 1\end{bmatrix},\quad \boldsymbol{x}_0=\begin{bmatrix}2.9\\ 2.6\end{bmatrix}$$

求解初值问题并画出在 $t\geqslant 0$ 时粒子的轨迹.

解 求出 A 的特征值为 $\lambda_1=6$ 和 $\lambda_2=-1$，相应的特征向量是

$$\boldsymbol{v}_1=(-5,2)\quad 和\quad \boldsymbol{v}_2=(1,1)$$

对任意的常数 c_1 和 c_2，函数

$$\boldsymbol{x}(t)=c_1\boldsymbol{v}_1\mathrm{e}^{\lambda_1 t}+c_2\boldsymbol{v}_2\mathrm{e}^{\lambda_2 t}=c_1\begin{bmatrix}-5\\ 2\end{bmatrix}\mathrm{e}^{6t}+c_2\begin{bmatrix}1\\ 1\end{bmatrix}\mathrm{e}^{-t}$$

是 $\boldsymbol{x}'=A\boldsymbol{x}$ 的解，我们要求 c_1 和 c_2 满足 $\boldsymbol{x}(0)=\boldsymbol{x}_0$，即

$$c_1\begin{bmatrix}-5\\ 2\end{bmatrix}+c_2\begin{bmatrix}1\\ 1\end{bmatrix}=\begin{bmatrix}2.9\\ 2.6\end{bmatrix}\quad 或\quad \begin{bmatrix}-5 & 1\\ 2 & 1\end{bmatrix}\begin{bmatrix}c_1\\ c_2\end{bmatrix}=\begin{bmatrix}2.9\\ 2.6\end{bmatrix}$$

求得 $c_1=-3/70$ 和 $c_2=188/70$，所求的函数是

$$\boldsymbol{x}(t)=\frac{-3}{70}\begin{bmatrix}-5\\ 2\end{bmatrix}\mathrm{e}^{6t}+\frac{188}{70}\begin{bmatrix}1\\ 1\end{bmatrix}\mathrm{e}^{-t}$$

$\boldsymbol{x}$ 和其他解的轨迹如图 5-23 所示.

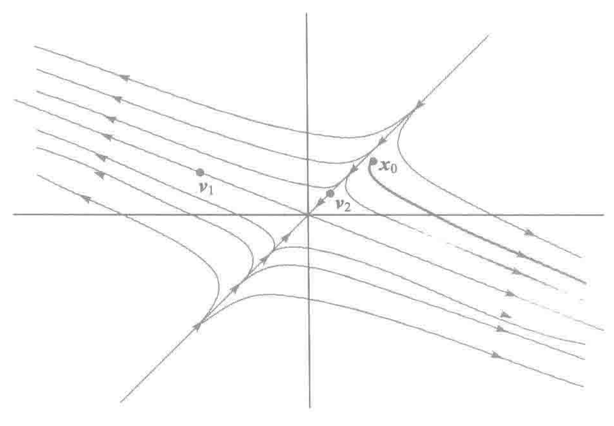


图 5-23 原点是鞍点 ■

在图 5-23 中，原点称为动力系统的**鞍点**，因为有些轨迹开始时趋近原点，然后又改变方向远离原点而去. 当矩阵 A 既有正的特征值，又有负的特征值时，就会出现鞍点. 最大排斥方向在过原点和 $\boldsymbol{v}_1$ 的直线上，对应于正的特征值，最大的吸引方向在过原点和 $\boldsymbol{v}_2$ 的直线上，对应负的特征值.

解耦动力系统

在下面的讨论中将看到，当 $n\times n$ 矩阵 A 有 n 个线性无关的特征向量，即 A 可对角化时，例 1 和例 2 所用的方法能够用来产生由 $\boldsymbol{x}'=A\boldsymbol{x}$ 所描述的动力系统的基础解系.

设 A 的特征函数是

$$\boldsymbol{v}_1\mathrm{e}^{\lambda_1 t},\cdots,\boldsymbol{v}_n\mathrm{e}^{\lambda_n t}$$

$\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$ 是线性无关的特征向量. 令 $P=[\boldsymbol{v}_1 \quad \cdots \quad \boldsymbol{v}_n]$，$D$ 是主对角线元素为 $\lambda_1,\cdots,\lambda_n$ 的对角矩阵，因此有 $A=PDP^{-1}$，现做变量代换，由

$$\boldsymbol{y}(t)=P^{-1}\boldsymbol{x}(t) \quad 或 \quad \boldsymbol{x}(t)=P\boldsymbol{y}(t)$$

定义一个新的函数 $\boldsymbol{y}$，代入 $\boldsymbol{x}'=A\boldsymbol{x}$，有

$$\frac{\mathrm{d}}{\mathrm{d}t}(P\boldsymbol{y})=A(P\boldsymbol{y})=(PDP^{-1})P\boldsymbol{y}=PD\boldsymbol{y} \tag{5}$$

由于 P 是常数矩阵，（5）的左边是 $P\boldsymbol{y}'$，在（5）的两边左乘 P^{-1}，得 $\boldsymbol{y}'=D\boldsymbol{y}$ 或

$$\begin{bmatrix} y_1'(t) \\ y_2'(t) \\ \vdots \\ y_n'(t) \end{bmatrix}=\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}\begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{bmatrix}$$

因为数值函数 y_k 的导数 $y_i'(t)$ 仅依赖于 y_k，故从 $\boldsymbol{x}$ 到 $\boldsymbol{y}$ 的变量代换解耦了微分方程组.（回忆 5.6 节类似的变量代换.）由 $y_1'=\lambda_1 y_1$，有 $y_1(t)=c_1\mathrm{e}^{\lambda_1 t}$，对 $y_2,\cdots,y_n$ 也有类似的公式. 因此

$$\boldsymbol{y}(t)=\begin{bmatrix} c_1\mathrm{e}^{\lambda_1 t} \\ \vdots \\ c_n\mathrm{e}^{\lambda_n t} \end{bmatrix}，其中\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}=\boldsymbol{y}(0)=P^{-1}\boldsymbol{x}(0)=P^{-1}\boldsymbol{x}_0$$

为了得到原方程组的通解 $\boldsymbol{x}$，计算

$$\begin{aligned}\boldsymbol{x}(t)&=P\boldsymbol{y}(t)=[\boldsymbol{v}_1 \quad \cdots \quad \boldsymbol{v}_n]\,\boldsymbol{y}(t)\\&=c_1\boldsymbol{v}_1\mathrm{e}^{\lambda_1 t}+\cdots+c_n\boldsymbol{v}_n\mathrm{e}^{\lambda_n t}\end{aligned}$$

这就是在例 1 求得的特征函数的扩展.

复特征值

在例 3 中，实矩阵 A 有共轭复特征值 λ 和 $\bar{\lambda}$，对应的特征向量为 $\boldsymbol{v}$ 和 $\bar{\boldsymbol{v}}$（回顾 5.5 节，实矩阵的复特征值和特征向量共轭出现）. 因此 $\boldsymbol{x}'=A\boldsymbol{x}$ 的两个解是

$$\boldsymbol{x}_1(t)=\boldsymbol{v}\mathrm{e}^{\lambda t} \quad 和 \quad \boldsymbol{x}_2(t)=\bar{\boldsymbol{v}}\mathrm{e}^{\bar{\lambda}t} \tag{6}$$

虽然这些复特征函数对某些计算（尤其是电子工程）是方便的，但对多数的实际应用而言，实函数要更合适一些.所幸的是，$\boldsymbol{x}_1$ 的实部和虚部是 $\boldsymbol{x}'=A\boldsymbol{x}$ 的（实数）解，因为它们是形如（6）的解的线性组合：

$$\mathrm{Re}(\boldsymbol{v}\mathrm{e}^{\lambda t})=\frac{1}{2}[\boldsymbol{x}_1(t)+\overline{\boldsymbol{x}_1(t)}]，\quad \mathrm{Im}(\boldsymbol{v}\mathrm{e}^{\lambda t})=\frac{1}{2\mathrm{i}}[\boldsymbol{x}_1(t)-\overline{\boldsymbol{x}_1(t)}]$$

为理解 $\mathrm{Re}(\boldsymbol{v}\mathrm{e}^{\lambda t})$ 的本质，回忆在微积分中，对任意的数 x，指数函数 e^x 可以通过幂级数计算：

$$\mathrm{e}^x=1+x+\frac{1}{2!}x^2+\cdots+\frac{1}{n!}x^n+\cdots$$

当 λ 为复数时，可利用该级数定义 $\mathrm{e}^{\lambda t}$：

$$\mathrm{e}^{\lambda t}=1+(\lambda t)+\frac{1}{2!}(\lambda t)^2+\cdots+\frac{1}{n!}(\lambda t)^n+\cdots$$

记 $\lambda=a+b\mathrm{i}$（a 和 b 是实数），对余弦和正弦函数利用相似的幂级数，可以证明

$$e^{(a+bi)t}=e^{at}\cdot e^{ibt}=e^{at}(\cos bt+i\sin bt) \tag{7}$$

因此

$$\begin{aligned}\boldsymbol{v}e^{\lambda t}&=(\operatorname{Re}\boldsymbol{v}+i\operatorname{Im}\boldsymbol{v})\cdot e^{at}(\cos bt+i\sin bt)\\&=[(\operatorname{Re}\boldsymbol{v})\cos bt-(\operatorname{Im}\boldsymbol{v})\sin bt]e^{at}\\&\quad+i[(\operatorname{Re}\boldsymbol{v})\sin bt+(\operatorname{Im}\boldsymbol{v})\cos bt]e^{at}\end{aligned}$$

故 $\boldsymbol{x}'=A\boldsymbol{x}$ 的两个实解是

$$\boldsymbol{y}_1(t)=\operatorname{Re}\boldsymbol{x}_1(t)=[(\operatorname{Re}\boldsymbol{v})\cos bt-(\operatorname{Im}\boldsymbol{v})\sin bt]e^{at}$$
$$\boldsymbol{y}_2(t)=\operatorname{Im}\boldsymbol{x}_1(t)=[(\operatorname{Re}\boldsymbol{v})\sin bt+(\operatorname{Im}\boldsymbol{v})\cos bt]e^{at}$$

可以证明，函数 $\boldsymbol{y}_1$ 和 $\boldsymbol{y}_2$ 是线性无关的函数（当 $b\neq 0$ 时）.[⊖]

例 3 图 5-24 的电路可以用下面的方程描述：

$$\begin{bmatrix}i_L'\\v_C'\end{bmatrix}=\begin{bmatrix}-R_2/L & -1/L\\1/C & -1/(R_1C)\end{bmatrix}\begin{bmatrix}i_L\\v_C\end{bmatrix}$$

其中 i_L 是通过电感 L 的电流，v_C 是电容器 C 的电压，设 R_1 为 5 欧姆，R_2 为 0.8 欧姆，C 为 0.1 法拉，L 为 0.4 亨利，若通过电感的初始电流为 3 安培，C 的初始电压为 3 伏，求计算 i_L 和 v_C 的公式.

图 5-24

解 由所给的数据，得 $A=\begin{bmatrix}-2 & -2.5\\10 & -2\end{bmatrix}$ 和 $\boldsymbol{x}_0=\begin{bmatrix}3\\3\end{bmatrix}$. 用 5.5 节的方法求得特征值 $\lambda=-2+5i$ 及对应的特征向量 $\boldsymbol{v}_1=\begin{bmatrix}i\\2\end{bmatrix}$. $\boldsymbol{x}'=A\boldsymbol{x}$ 的复数解是

$$\boldsymbol{x}_1(t)=\begin{bmatrix}i\\2\end{bmatrix}e^{(-2+5i)t} \text{ 和 } \boldsymbol{x}_2(t)=\begin{bmatrix}-i\\2\end{bmatrix}e^{(-2-5i)t}$$

的复线性组合，利用（7）式得

$$\boldsymbol{x}_1(t)=\begin{bmatrix}i\\2\end{bmatrix}e^{-2t}(\cos 5t+i\sin 5t)$$

取 $\boldsymbol{x}_1(t)$ 的实部和虚部，可以得到实数解：

$$\boldsymbol{y}_1(t)=\begin{bmatrix}-\sin 5t\\2\cos 5t\end{bmatrix}e^{-2t},\ \boldsymbol{y}_2(t)=\begin{bmatrix}\cos 5t\\2\sin 5t\end{bmatrix}e^{-2t}$$

因为 $\boldsymbol{y}_1$ 和 $\boldsymbol{y}_2$ 是线性无关的函数，故它们形成 $\boldsymbol{x}'=A\boldsymbol{x}$ 的二维实解向量空间的基. 因此，通解是

$$\boldsymbol{x}(t)=c_1\begin{bmatrix}-\sin 5t\\2\cos 5t\end{bmatrix}e^{-2t}+c_2\begin{bmatrix}\cos 5t\\2\sin 5t\end{bmatrix}e^{-2t}$$

为满足 $\boldsymbol{x}(0)=\begin{bmatrix}3\\3\end{bmatrix}$，要求 $c_1\begin{bmatrix}0\\2\end{bmatrix}+c_2\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}3\\3\end{bmatrix}$，解得 $c_1=1.5$ 和 $c_2=3$. 因此

[⊖] 由于 $\boldsymbol{x}_2(t)$ 是 $\boldsymbol{x}_1(t)$ 的复共轭，因此，它的实部和虚部分别是 $\boldsymbol{y}_1(t)$ 和 $-\boldsymbol{y}_2(t)$.因此，可以利用 $\boldsymbol{x}_1(t)$ 或 $\boldsymbol{x}_2(t)$，但不是同时使用两者，来产生 $\boldsymbol{x}'=A\boldsymbol{x}$ 的线性无关的解.

$$\boldsymbol{x}(t)=1.5\begin{bmatrix}-\sin 5t\\2\cos 5t\end{bmatrix}\mathrm{e}^{-2t}+3\begin{bmatrix}\cos 5t\\2\sin 5t\end{bmatrix}\mathrm{e}^{-2t}$$

或者

$$\begin{bmatrix}i_L(t)\\v_C(t)\end{bmatrix}=\begin{bmatrix}-1.5\sin 5t+3\cos 5t\\3\cos 5t+6\sin 5t\end{bmatrix}\mathrm{e}^{-2t}$$

见图 5-25.

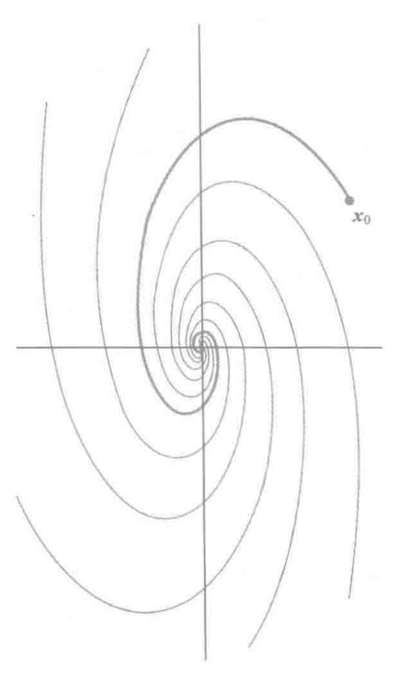


图 5-25　原点是螺线极点 ■

在图 5-25 中，原点称为动力系统的**螺线极点**. 旋转是由复特征值产生的正弦和余弦函数产生的. 因为系数 e^{-2t} 趋于零，故轨迹往里旋转. 注意，–2 是例 3 中的特征值的实部. 当 A 的复特征值的实部为正数时，轨迹往外旋转. 当特征值的实部为零时，轨迹形成绕原点的椭圆.

练习题

3×3 实矩阵 A 有特征值 -0.5 ,$0.2+0.3\mathrm{i}$ 和 $0.2-0.3\mathrm{i}$ ，对应的特征向量为

$$\boldsymbol{v}_1=\begin{bmatrix}1\\-2\\1\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}1+2\mathrm{i}\\4\mathrm{i}\\2\end{bmatrix},\quad \boldsymbol{v}_3=\begin{bmatrix}1-2\mathrm{i}\\-4\mathrm{i}\\2\end{bmatrix}$$

1. 利用复矩阵， A 能否对角化为 D ，即 $A=PDP^{-1}$.
2. 用复矩阵函数写出 $\boldsymbol{x}'=A\boldsymbol{x}$ 的通解，然后再求出实的通解.
3. 描述典型的轨迹形状.

习题 5.7

1. 在平面力场运动的粒子的位置向量满足 $x' = Ax$ .2×2 矩阵 A 有特征值 4 和 2，对应的特征向量为 $v_1 = \begin{bmatrix} -3 \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.假设 $x(0) = \begin{bmatrix} -6 \\ 1 \end{bmatrix}$，求粒子在时刻 t 的位置.
2. 设 2×2 矩阵 A 的特征值为−3 和−1，对应的特征向量为 $v_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$， $v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$， $x(t)$ 是粒子在时间 t 的位置. 求初值问题 $x' = Ax$， $x(0) = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

在习题 3~6 中，对 $t \geqslant 0, x(0) = (3, 2)$，求解初值问题 $x'(t) = Ax(t)$. 把原点分类为由 $x' = Ax$ 描述的动力系统的吸引子、排斥子或鞍点. 并求最大的吸引方向或排斥方向.当原点是鞍点时，画出典型的轨迹.

3. $A = \begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix}$
4. $A = \begin{bmatrix} -2 & -5 \\ 1 & 4 \end{bmatrix}$
5. $A = \begin{bmatrix} 7 & -1 \\ 3 & 3 \end{bmatrix}$
6. $A = \begin{bmatrix} 1 & -2 \\ 3 & -4 \end{bmatrix}$

在习题 7 和习题 8 中，做变量代换解耦方程 $x' = Ax$. 求出 P 和 D，写出方程 $x(t) = Py(t)$，并写出得到分离系统 $y' = Dy$ 的计算过程.

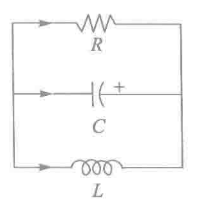
7. 习题 5 的矩阵 A .
8. 习题 6 的矩阵 A .

在习题 9~18 中，求 $x' = Ax$ 的包含复特征函数的通解，然后求得实通解，并描述典型轨迹的形状.

9. $A = \begin{bmatrix} -3 & 2 \\ -1 & -1 \end{bmatrix}$
10. $A = \begin{bmatrix} 3 & 1 \\ -2 & 1 \end{bmatrix}$
11. $A = \begin{bmatrix} -3 & -9 \\ 2 & 3 \end{bmatrix}$
12. $A = \begin{bmatrix} -7 & 10 \\ -4 & 5 \end{bmatrix}$
13. $A = \begin{bmatrix} 4 & -3 \\ 6 & -2 \end{bmatrix}$
14. $A = \begin{bmatrix} -2 & 1 \\ -8 & 2 \end{bmatrix}$
15. [M] $A = \begin{bmatrix} -8 & -12 & -6 \\ 2 & 1 & 2 \\ 7 & 12 & 5 \end{bmatrix}$
16. [M] $A = \begin{bmatrix} -6 & -11 & 16 \\ 2 & 5 & -4 \\ -4 & -5 & 10 \end{bmatrix}$
17. [M] $A = \begin{bmatrix} 30 & 64 & 23 \\ -11 & -23 & -9 \\ 6 & 15 & 4 \end{bmatrix}$
18. [M] $A = \begin{bmatrix} 53 & -30 & -2 \\ 90 & -52 & -3 \\ 20 & -10 & 2 \end{bmatrix}$
19. [M]求例 1 中电路的电压 v_1 和 v_2（作为时间 t 的函数）的公式，设 $R_1 = 1/5$ 欧姆，$R_2 = 1/3$ 欧姆，$C_1 = 4$ 法拉，$C_2 = 3$ 法拉，每个电容器的初始电压为 4 伏.
20. [M]求例 1 中电路的电压 v_1 和 v_2 的公式，设 $R_1 = 1/15$ 欧姆， $R_2 = 1/3$ 欧姆， $C_1 = 9$ 法拉，$C_2 = 2$ 法拉，每个电容器的初始电压为 3 伏.
21. [M]求例 3 中电路的电流 i_L 和电压 v_C 的公式，设 $R_1 = 1$ 欧姆， $R_2 = 0.125$ 欧姆， $C = 0.2$ 法拉，$L = 0.125$ 亨利，初始电流为 0 安培，初始电压为 15 伏.
22. [M]例 6 的电路由下列方程描述：

$$\begin{bmatrix} i'_L \\ v'_C \end{bmatrix} = \begin{bmatrix} 0 & 1/L \\ -1/C & -1/(RC) \end{bmatrix} \begin{bmatrix} i_L \\ v_C \end{bmatrix}$$

其中，i_L 是通过电感 L 的电流，v_C 是电容器 C 的电压. 设 $R = 0.5$ 欧姆，$C = 2.5$ 法拉，$L = 0.5$ 亨利，初始电流为 0 安培，初始电压为 12 伏，求计算 i_L 和 v_C 的公式.



练习题答案

1. A 可对角化，因为 A 有三个不同的特征值. 当使用复数时，5.1 节的定理 2 以及 5.3 节的定理 5 仍然有效（证明基本上与实数相同）.
2. 通解具有以下形式

$$\boldsymbol{x}(t)=c_1\begin{bmatrix}1\\-2\\1\end{bmatrix}\mathrm{e}^{-0.5t}+c_2\begin{bmatrix}1+2\mathrm{i}\\4\mathrm{i}\\2\end{bmatrix}\mathrm{e}^{(0.2+0.3\mathrm{i})t}+c_3\begin{bmatrix}1-2\mathrm{i}\\-4\mathrm{i}\\2\end{bmatrix}\mathrm{e}^{(0.2-03\mathrm{i})t}$$

其中 c_1,c_2,c_3 为任意复数. $\boldsymbol{x}(t)$ 的第一项是实的，其他两个实解可以用 $\boldsymbol{x}(t)$ 的第二项：

$$\begin{bmatrix}1+2\mathrm{i}\\4\mathrm{i}\\2\end{bmatrix}\mathrm{e}^{0.2t}(\cos 0.3t+\mathrm{i}\sin 0.3t)$$

的实部和虚部来产生. 实通解为

$$c_1\begin{bmatrix}1\\-2\\1\end{bmatrix}\mathrm{e}^{-0.5t}+c_2\begin{bmatrix}\cos 0.3t-2\sin 0.3t\\-4\sin 0.3t\\2\cos 0.3t\end{bmatrix}\mathrm{e}^{0.2t}+c_3\begin{bmatrix}\sin 0.3t+2\cos 0.3t\\4\cos 0.3t\\2\sin 0.3t\end{bmatrix}\mathrm{e}^{0.2t}$$

其中 c_1,c_2,c_3 为实数.

3. 因为有负的指数因子，当 $c_2=c_3=0$ 时的任一解逼近原点，其他的解有无界的分量，轨迹朝外旋转. 小心不要将这个问题错当成 5.6 节的问题. 在 5.6 节中，逼近原点的条件是特征值的绝对值小于 1，使得 $|\lambda|^k\to 0$. 这里的条件是特征值的实部必须是负的，才使得 $\mathrm{e}^{\lambda t}\to 0$.

5.8 特征值的迭代估计

在线性代数的科学应用中，很少能精确知道特征值. 所幸的是，一个较精确的数值近似通常能达到满意的效果. 实际上，某些应用只需要粗略估计最大特征值. 下面介绍的第 1 个算法很适合这种情况. 同样，它为快速估计其他特征值的更有效的方法提供了基础.

幂算法

幂算法适用于 $n\times n$ 矩阵 A 有**严格占优特征值**（亦称主特征值）λ_1 的情况. λ_1 为主特征值的意思是 λ_1 的绝对值比其他特征值的绝对值都大. 此时，幂算法产生一个近似 λ_1 的数列和一个近似主特征向量的向量序列. 此方法的背景来源于 5.6 节开头的特征向量分解.

为简单起见，假设 A 可对角化，特征向量 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$ 是 $\mathbb{R}^n$ 的基，并且 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$ 已经过排列，使对应的特征值 $\lambda_1,\cdots,\lambda_n$ 的绝对值递减，λ_1 是主特征值，即有

$$|\lambda_1|>|\lambda_2|\geqslant|\lambda_3|\geqslant\cdots\geqslant|\lambda_n| \quad (1)$$

（$>$：严格大）

就像我们在 5.6 节式（2）中看到的，假设 $x\in\mathbb{R}^n, x=c_1\boldsymbol{v}_1+\cdots+c_n\boldsymbol{v}_n$，那么

$$A^k\boldsymbol{x}=c_1(\lambda_1)^k\boldsymbol{v}_1+c_2(\lambda_2)^k\boldsymbol{v}_2+\cdots+c_n(\lambda_n)^k\boldsymbol{v}_n\quad(k=1,2,\cdots)$$

假设 $c_1 \neq 0$，等式除于 $(\lambda_1)^k$，

$$\frac{1}{(\lambda_1)^k} A^k \boldsymbol{x} = c_1\boldsymbol{v}_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k \boldsymbol{v}_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k \boldsymbol{v}_n \quad (k=1,2,\cdots) \tag{2}$$

由（1），所有分数 $\lambda_2/\lambda_1,\cdots,\lambda_n/\lambda_1$ 的值都小于 1，因此，它们的幂趋于零. 故

$$\text{当 } k\to\infty \text{ 时，} (\lambda_1)^{-k} A^k \boldsymbol{x} \to c_1\boldsymbol{v}_1 \tag{3}$$

因此对足够大的 k，$A^k\boldsymbol{x}$ 的数量倍的方向几乎与特征向量 $c_1\boldsymbol{v}_1$ 的方向相同. 由于正的数量倍不会改变向量的方向，若给定 $c_1 \neq 0$，$A^k\boldsymbol{x}$ 的方向几乎与 $\boldsymbol{v}_1$ 或 $-\boldsymbol{v}_1$ 一致.

例 1 设 $A=\begin{bmatrix}1.8 & 0.8\\0.2 & 1.2\end{bmatrix}, \boldsymbol{v}_1=\begin{bmatrix}4\\1\end{bmatrix}, \boldsymbol{x}=\begin{bmatrix}-0.5\\1\end{bmatrix}$，那么 A 的特征值是 2 和 1，$\lambda_1=2$ 的特征空间是过原点和 $\boldsymbol{v}_1$ 的直线. 对 $k=0,\cdots,8$，计算 $A^k\boldsymbol{x}$ 并画出过原点和 $A^k\boldsymbol{x}$ 的直线. 当 k 增大时会出现什么情况？

解 开头三项的计算是

$$A\boldsymbol{x}=\begin{bmatrix}1.8 & 0.8\\0.2 & 1.2\end{bmatrix}\begin{bmatrix}-0.5\\1\end{bmatrix}=\begin{bmatrix}-0.1\\1.1\end{bmatrix}$$

$$A^2\boldsymbol{x}=A(A\boldsymbol{x})=\begin{bmatrix}1.8 & 0.8\\0.2 & 1.2\end{bmatrix}\begin{bmatrix}-0.1\\1.1\end{bmatrix}=\begin{bmatrix}0.7\\1.3\end{bmatrix}$$

$$A^3\boldsymbol{x}=A(A^2\boldsymbol{x})=\begin{bmatrix}1.8 & 0.8\\0.2 & 1.2\end{bmatrix}\begin{bmatrix}0.7\\1.3\end{bmatrix}=\begin{bmatrix}2.3\\1.7\end{bmatrix}$$

其他项的计算在表 5-1 中给出.

表 5-1 向量的迭代

k	0	1	2	3	4	5	6	7	8
$A^k\boldsymbol{x}$	$\begin{bmatrix}-0.5\\1\end{bmatrix}$	$\begin{bmatrix}-0.1\\1.1\end{bmatrix}$	$\begin{bmatrix}0.7\\1.3\end{bmatrix}$	$\begin{bmatrix}2.3\\1.7\end{bmatrix}$	$\begin{bmatrix}5.5\\2.5\end{bmatrix}$	$\begin{bmatrix}11.9\\4.1\end{bmatrix}$	$\begin{bmatrix}24.7\\7.3\end{bmatrix}$	$\begin{bmatrix}50.3\\13.7\end{bmatrix}$	$\begin{bmatrix}101.5\\26.5\end{bmatrix}$

图 5-26 显示了向量 $\boldsymbol{x}, A\boldsymbol{x}, A\boldsymbol{x}^2, A\boldsymbol{x}^3, A\boldsymbol{x}^4$. 其余向量太长难于显示，但画出的线段显示出这些向量的方向. 事实上，我们真正想看到的是向量的方向而不是向量本身. 这些直线看起来是逼近表示 $\boldsymbol{v}_1$ 生成的特征空间的直线. 更加明确地，由 $A^k\boldsymbol{x}$ 确定的直线（子空间）与由 $\boldsymbol{v}_1$ 确定的直线（特征空间）之间的夹角（当 $k\to\infty$ 时）趋于零.

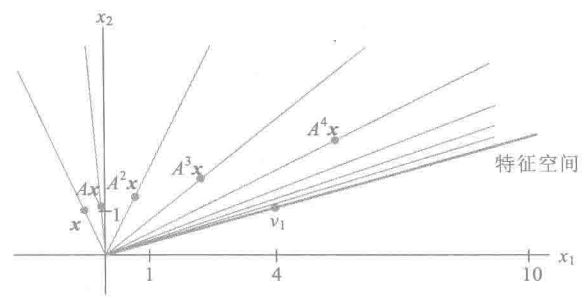


图 5-26 由 $\boldsymbol{x}, A\boldsymbol{x}, A^2\boldsymbol{x},\cdots, A^7\boldsymbol{x}$ 确定的方向 ■

当$c_1 \neq 0$时，我们可以对（3）中的向量$(\lambda_1)^{-k} A^k \boldsymbol{x}$倍乘，使它们收敛于$c_1\boldsymbol{v}_1$，但不能对$A^k\boldsymbol{x}$作这种倍乘，因为我们并不知道$\lambda_1$. 不过我们可以倍乘$A^k\boldsymbol{x}$，使它的最大分量为 1. 这样，所得序列$\{\boldsymbol{x}_k\}$将收敛于$\boldsymbol{v}_1$的倍数，$\boldsymbol{v}_1$的倍数的最大分量也是 1. 图 5-27 显示了例 1 的倍乘序列. 我们还可以通过序列$\{\boldsymbol{x}_k\}$来估值λ_1. 当$\boldsymbol{x}_k$接近于λ_1的特征向量时，向量$A\boldsymbol{x}_k$也接近于$\lambda_1\boldsymbol{x}_k$，即$A\boldsymbol{x}_k$的每一分量接近于$\lambda_1$与$\boldsymbol{x}_k$相应分量的乘积.因为$\boldsymbol{x}_k$的最大分量是 1，故$A\boldsymbol{x}_k$的最大分量接近于$\lambda_1$.（这些结论的证明省略了.）

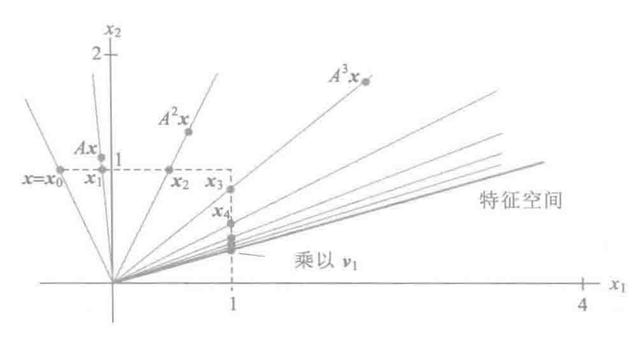


图 5-27 $\boldsymbol{x}, A\boldsymbol{x}, A^2\boldsymbol{x}, \cdots, A^7\boldsymbol{x}$的倍乘

> **估计严格占优特征值的幂算法**
>
> 1. 选择一个最大分量为 1 的初始向量$\boldsymbol{x}_0$.
> 2. 对$k = 0, 1, \cdots$,
>   a. 计算$A\boldsymbol{x}_k$.
>   b. 设μ_k是$A\boldsymbol{x}_k$中绝对值最大的一个分量.
>   c. 计算$\boldsymbol{x}_{k+1} = (1/\mu_k)A\boldsymbol{x}_k$.
> 3. 几乎对所有选择的$\boldsymbol{x}_0$，序列$\{\mu_k\}$近似于主特征值，而序列$\{\boldsymbol{x}_k\}$近似于对应的特征向量.

例 2 设$A = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}, \boldsymbol{x}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. 应用幂算法求$A$的主特征值和对应特征向量的近似值，计算到$k = 5$.

解 本例及下一个例子的计算用 MATLAB 完成，虽然这里看到的有效数字较少，但计算精确到 16 位. 先计算$A\boldsymbol{x}_0$，然后将$A\boldsymbol{x}_0$的最大分量μ_0单位化：$A\boldsymbol{x}_0 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$, $\mu_0 = 5$.

用$1/\mu_0$倍乘$A\boldsymbol{x}_0$得到$\boldsymbol{x}_1$，计算$A\boldsymbol{x}_1$，并将$A\boldsymbol{x}_1$的最大分量单位化：

$$\boldsymbol{x}_1 = \frac{1}{\mu_0}A\boldsymbol{x}_0 = \frac{1}{5}\begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.4 \end{bmatrix}$$

$$A\boldsymbol{x}_1 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0.4 \end{bmatrix} = \begin{bmatrix} 8 \\ 1.8 \end{bmatrix}, \ \mu_1 = 8$$

又用$1/\mu_1$倍乘$A\boldsymbol{x}_1$得到$\boldsymbol{x}_2$，计算$A\boldsymbol{x}_2$，将$A\boldsymbol{x}_2$的最大分量单位化：

$$x_2 = \frac{1}{\mu_1} A x_1 = \frac{1}{8}\begin{bmatrix} 8 \\ 1.8 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.225 \end{bmatrix}$$

$$A x_2 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0.225 \end{bmatrix} = \begin{bmatrix} 7.125 \\ 1.450 \end{bmatrix},\ \mu_2 = 7.125$$

又用 $1/\mu_2$ 倍乘 Ax_2 得到 x_3，一直重复做下去，用 MATLAB 计算的前 5 次迭代结果列在表 5-2 中.

表 5-2　例 2 的幂算法

k	0	1	2	3	4	5
x_k	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.4 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.225 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.2035 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.2005 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.20007 \end{bmatrix}$
Ax_k	$\begin{bmatrix} 5 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 8 \\ 1.8 \end{bmatrix}$	$\begin{bmatrix} 7.125 \\ 1.450 \end{bmatrix}$	$\begin{bmatrix} 7.0175 \\ 1.4070 \end{bmatrix}$	$\begin{bmatrix} 7.0025 \\ 1.4010 \end{bmatrix}$	$\begin{bmatrix} 7.00036 \\ 1.40014 \end{bmatrix}$
μ_k	5	8	7.125	7.0175	7.0025	7.00036

从表 5-2 的数据明显看出，$\{x_k\}$ 接近于（1,0.2），$\{\mu_k\}$ 接近于 7. 这样，（1,0.2）就是特征向量，7 是主特征值. 容易通过计算来验证 $A\begin{bmatrix} 1 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 7 \\ 1.4 \end{bmatrix} = 7\begin{bmatrix} 1 \\ 0.2 \end{bmatrix}$. ■

例 2 的序列 $\{\mu_k\}$ 很快收敛于 $\lambda_1 = 7$ 是因为 A 的第 2 个特征值 λ_2 要比 λ_1 小得多（事实上，$\lambda_2 = 1$）. 通常收敛的快慢取决于比率 $|\lambda_2/\lambda_1|$，因为用 $A^k x$ 的倍数来估值 $c_1 v_1$，主要误差来源于（2）中的向量 $c_2(\lambda_2/\lambda_1)^k v_2$（其他的分数 λ_j/λ_1 可能更小）. 若 $|\lambda_2/\lambda_1|$ 接近于 1，则 $\{\mu_k\}$ 和 $\{x_k\}$ 会收敛得很慢，此时，可能需要选择其他近似的方法.

不过幂算法还存在这样的很小可能性，即随机选择的初始向量 x 在 v_1 的方向上没有分量（当 $c_1 = 0$ 时），但计算机在计算 x_k 时所产生的舍入误差有可能产生一个向量，该向量在 v_1 的方向上至少有一个小分量. 如果这样的话，x_k 将开始收敛于数倍 v_1 的某个向量.

逆幂法

在知道特征值 λ 的一个较好的初步估值 α 后，逆幂法可用来对任一特征值近似估值. 此时，我们令 $B = (A - \alpha I)^{-1}$，并对 B 应用幂算法. 可以证明，若 $\lambda_1, \cdots, \lambda_n$ 是 A 的特征值，则 B 的特征值是

$$\frac{1}{\lambda_1 - \alpha}, \frac{1}{\lambda_2 - \alpha}, \cdots, \frac{1}{\lambda_n - \alpha}$$

而且 A 的对应于 $\lambda_1, \cdots, \lambda_n$ 的特征向量是 B 的对应于上面这些特征值的特征向量.（见习题 15 和 16.）

例如，假设 α 更接近 λ_2 而不是其他特征值，那么 $1/(\lambda_2 - \alpha)$ 将是 B 的严格主特征值. 假如 α 确实接近于 λ_2，那么 $1/(\lambda_2 - \alpha)$ 比 B 的其他特征值要"大"得多. 几乎对所有选择的 x_0，逆幂法会快速逼近 λ_2，下列算法给出了详细步骤.

估计 A 的特征值 λ 的逆幂法

1. 选择一个非常接近于 λ 的初始值 α.
2. 选择一个最大分量为 1 的初始向量 x_0.

> 3. 对 $k=0,1,\cdots$,
>   a. 从 $(A-\alpha I)\boldsymbol{y}_k=\boldsymbol{x}_k$ 解出 $\boldsymbol{y}_k$.
>   b. 设 μ_k 是 $\boldsymbol{y}_k$ 绝对值最大的分量.
>   c. 计算 $v_k=\alpha+(1/\mu_k)$.
>   d. 计算 $\boldsymbol{x}_{k+1}=(1/\mu_k)\boldsymbol{y}_k$.
>
> 4. 几乎对所有选择的 $\boldsymbol{x}_0$，序列 $\{\boldsymbol{v}_k\}$ 趋向于 A 的特征值 λ，而序列 $\{\boldsymbol{x}_k\}$ 趋向于对应的特征向量.

注意，B 或者是 $(A-\alpha I)^{-1}$ 并没有出现在算法中. 在求序列的下一个向量时并没有计算 $(A-\alpha I)^{-1}\boldsymbol{x}_k$，而是通过解方程 $(A-\alpha I)\boldsymbol{y}_k=\boldsymbol{x}_k$ 来得到 $\boldsymbol{y}_k$ 的（然后用倍乘 $\boldsymbol{y}_k$ 来产生 $\boldsymbol{x}_{k+1}$）. 因为对每个 k 一定可以从该方程解出 $\boldsymbol{y}_k$，故 $A-\alpha I$ 的 LU 分解式将加快计算过程.

例 3　在某些应用中，通常需要知道 A 的最小特征值，也需要对该特征值作粗略的估计. 设 2.1，3.3 和 1.9 是下列矩阵 A 的特征值的估值. 求最小特征值，小数位精确到 6 位.

$$A=\begin{bmatrix}10 & -8 & -4\\ -8 & 13 & 4\\ -4 & 5 & 4\end{bmatrix}$$

解　两个最小特征值看起来很接近，因此我们对 $A-1.9I$ 应用逆幂法. MATLAB 的计算结果列在表 5-3 中. 这里的 $\boldsymbol{x}_0$ 随机选取，$\boldsymbol{y}_k=(A-1.9I)^{-1}\boldsymbol{x}_k$, μ_k 是 $\boldsymbol{y}_k$ 的最大分量，$v_k=1.9+1/\mu_k$, $\boldsymbol{x}_{k+1}=(1/\mu_k)\boldsymbol{y}_k$. 可以看到，初始的特征值估计得非常好，逆幂法产生的序列收敛很快，准确的特征值是 2. ■

表 5-3　逆幂法

k	0	1	2	3	4
$\boldsymbol{x}_k$	$\begin{bmatrix}1\\1\\1\end{bmatrix}$	$\begin{bmatrix}0.5736\\0.0646\\1\end{bmatrix}$	$\begin{bmatrix}0.5054\\0.0045\\1\end{bmatrix}$	$\begin{bmatrix}0.5004\\0.0003\\1\end{bmatrix}$	$\begin{bmatrix}0.50003\\0.00002\\1\end{bmatrix}$
$\boldsymbol{y}_k$	$\begin{bmatrix}4.45\\0.50\\7.76\end{bmatrix}$	$\begin{bmatrix}5.0131\\0.0442\\9.9197\end{bmatrix}$	$\begin{bmatrix}5.0012\\0.0031\\9.9949\end{bmatrix}$	$\begin{bmatrix}5.0001\\0.0002\\9.9996\end{bmatrix}$	$\begin{bmatrix}5.000006\\0.000015\\9.999975\end{bmatrix}$
μ_k	7.76	9.919 7	9.994 9	9.999 6	9.999 975
v_k	2.03	2.000 8	2.000 05	2.000 004	2.000 000 2

假如没有矩阵最小特征值的估值，可以简单取 $\alpha=0$ 来使用逆幂法. 假如最小特征值更接近于零而不是其他的特征值，那么这种取值是合理的.

对很多简单的情况，本节给出的两个算法是实用的，它们为特征值估值问题提供了入门知识. 另一个更全面和广泛使用的迭代算法是 QR 算法. 比如，MATLAB 的命令 eig(A)的核心是 QR 算法，这个命令能快速计算出 A 的特征值和特征向量. 在 5.2 节的习题中对 QR 算法有过简短的描述，进一步的细节参见有关现代数值分析的教材.

练习题

你怎样断定给出的向量 $\boldsymbol{x}$ 是矩阵 A 的某个特征向量的好的近似值；如果是，又怎样估计相应的特征值？用下面给出的矩阵 A 和 $\boldsymbol{x}$ 做练习.

$$A=\begin{bmatrix}5 & 8 & 4\\ 8 & 3 & -1\\ 4 & -1 & 2\end{bmatrix} \text{ 和 } \boldsymbol{x}=\begin{bmatrix}1.0\\ -4.3\\ 8.1\end{bmatrix}$$

习题 5.8

在习题 1~4 中，给出矩阵 A 和由幂算法产生的序列 $\{\boldsymbol{x}_k\}$. 利用这些数据对 A 的最大特征值进行估计，并求出对应的特征向量.

1. $A=\begin{bmatrix}4 & 3\\ 1 & 2\end{bmatrix};\begin{bmatrix}1\\ 0\end{bmatrix},\begin{bmatrix}1\\ 0.25\end{bmatrix},\begin{bmatrix}1\\ 0.315\,8\end{bmatrix},\begin{bmatrix}1\\ 0.329\,8\end{bmatrix},\begin{bmatrix}1\\ 0.332\,6\end{bmatrix}$

2. $A=\begin{bmatrix}1.8 & -0.8\\ -3.2 & 4.2\end{bmatrix};\begin{bmatrix}1\\ 0\end{bmatrix},\begin{bmatrix}-0.562\,5\\ 1\end{bmatrix},\begin{bmatrix}-0.302\,1\\ 1\end{bmatrix},\begin{bmatrix}-0.260\,1\\ 1\end{bmatrix},\begin{bmatrix}-0.252\,0\\ 1\end{bmatrix}$

3. $A=\begin{bmatrix}0.5 & 0.2\\ 0.4 & 0.7\end{bmatrix};\begin{bmatrix}1\\ 0\end{bmatrix},\begin{bmatrix}1\\ 0.8\end{bmatrix},\begin{bmatrix}0.687\,5\\ 1\end{bmatrix},\begin{bmatrix}0.557\,7\\ 1\end{bmatrix},\begin{bmatrix}0.518\,8\\ 1\end{bmatrix}$

4. $A=\begin{bmatrix}4.1 & -6\\ 3 & -4.4\end{bmatrix};\begin{bmatrix}1\\ 1\end{bmatrix},\begin{bmatrix}1\\ 0.736\,8\end{bmatrix},\begin{bmatrix}1\\ 0.754\,1\end{bmatrix},\begin{bmatrix}1\\ 0.749\,0\end{bmatrix},\begin{bmatrix}1\\ 0.750\,2\end{bmatrix}$

5. 设 $A=\begin{bmatrix}15 & 16\\ -20 & -21\end{bmatrix}$，向量 $\boldsymbol{x},\cdots,A^5\boldsymbol{x}$ 分别是：

$$\begin{bmatrix}1\\ 1\end{bmatrix},\begin{bmatrix}31\\ -41\end{bmatrix},\begin{bmatrix}-191\\ 241\end{bmatrix},\begin{bmatrix}991\\ -1241\end{bmatrix},\begin{bmatrix}-4\,991\\ 6\,241\end{bmatrix},\begin{bmatrix}24\,991\\ -31\,241\end{bmatrix}$$

寻找一个第 2 个分量为 1 且接近于 A 的一个特征向量的向量，保留 4 位小数. 验证你的估值，并对 A 的主特征值进行估计.

6. 设 $A=\begin{bmatrix}-2 & -3\\ 6 & 7\end{bmatrix}$. 利用下列序列 $\boldsymbol{x},A\boldsymbol{x},\cdots,A^5\boldsymbol{x}$：

$$\begin{bmatrix}1\\ 1\end{bmatrix},\begin{bmatrix}-5\\ 13\end{bmatrix},\begin{bmatrix}-29\\ 61\end{bmatrix},\begin{bmatrix}-125\\ 253\end{bmatrix},\begin{bmatrix}-509\\ 1\,021\end{bmatrix},\begin{bmatrix}-2\,045\\ 4\,093\end{bmatrix}$$

重做习题 5.

[M] 习题 7~12 需要用到 MATLAB 或其他辅助的计算工具. 在习题 7 和 8 中，利用幂算法及给出的 $\boldsymbol{x}_0$，对 $k=1,\cdots,5$，算出 $\{\boldsymbol{x}_k\}$ 和 $\{\mu_k\}$. 在习题 9 和 10 中，算出 μ_5 和 μ_6.

7. $A=\begin{bmatrix}6 & 7\\ 8 & 5\end{bmatrix}$，$\boldsymbol{x}_0=\begin{bmatrix}1\\ 0\end{bmatrix}$

8. $A=\begin{bmatrix}2 & 1\\ 4 & 5\end{bmatrix}$，$\boldsymbol{x}_0=\begin{bmatrix}1\\ 0\end{bmatrix}$

9. $A=\begin{bmatrix}8 & 0 & 12\\ 1 & -2 & 1\\ 0 & 3 & 0\end{bmatrix}$，$\boldsymbol{x}_0=\begin{bmatrix}1\\ 0\\ 0\end{bmatrix}$

10. $A=\begin{bmatrix}1 & 2 & -2\\ 1 & 1 & 9\\ 0 & 1 & 9\end{bmatrix}$，$\boldsymbol{x}_0=\begin{bmatrix}1\\ 0\\ 0\end{bmatrix}$

如果知道某个特征向量的近似值，就可以对特征值作另一估值，若 $A\boldsymbol{x}=\lambda\boldsymbol{x}$，则 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}(\lambda\boldsymbol{x})=\lambda(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x})$，瑞利商 $R(\boldsymbol{x})=\dfrac{\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}$ 等于 λ. 假如 x 接近于 λ 对应的特征向量，那么瑞利商就接近于 λ. 当 A 为对称矩阵时 ($A^{\mathrm{T}}=A$)，瑞利商 $R(\boldsymbol{x}_k)=(\boldsymbol{x}_k^{\mathrm{T}}A\boldsymbol{x}_k)/\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{x}_k$ 的精确数字大致是幂算法产生的倍乘因子 μ_k 的 2 倍，在习题 11 和 12 中通过计算 μ_k 和 $R(\boldsymbol{x}_k)(k=1,\cdots,4)$ 来验证增加的精确度.

11. $A=\begin{bmatrix}5 & 2\\ 2 & 2\end{bmatrix}$，$\boldsymbol{x}_0=\begin{bmatrix}1\\ 0\end{bmatrix}$

12. $A=\begin{bmatrix}-3 & 2\\ 2 & 0\end{bmatrix}$，$\boldsymbol{x}_0=\begin{bmatrix}1\\ 0\end{bmatrix}$

习题 13 和 14 应用于一个 3×3 矩阵 A，其特征值被估计为 4，−4 和 3.

13. 对接近于 4 和-4 但绝对值不同的特征值，幂算法还有效吗？
14. 对接近于 4 和-4 但绝对值相同的特征值，描述怎样才能得到估计接近于4的特征值的序列.
15. 假设 $A\boldsymbol{x}=\lambda\boldsymbol{x}$，$\boldsymbol{x}\neq\boldsymbol{0}$，数 α 不是 A 的特征值，令 $B=(A-\alpha I)^{-1}$，在等式 $A\boldsymbol{x}=\lambda\boldsymbol{x}$ 的两边减去 $\alpha\boldsymbol{x}$，用代数方法证明 $1/(\lambda-\alpha)$ 是 B 的特征值，$\boldsymbol{x}$ 是其对应的特征向量.
16. 设 μ 是习题 15 矩阵 B 的特征值，$\boldsymbol{x}$ 是对应的特征向量，即有 $(A-\alpha I)^{-1}\boldsymbol{x}=\mu\boldsymbol{x}$，利用这个等式求 A 的用 μ 和 α 表示的特征值.（注意：因为 B 可逆，故 $\mu\neq 0$.）
17. [M]设 $\boldsymbol{x}_0=(1,0,0)$，利用逆幂法对例 3 中的矩阵 A 的中间特征值进行估值，计算精确到 4 位小数.
18. [M]设 A 为习题 9 给出的矩阵，$\boldsymbol{x}_0=(1,0,0)$，利用逆幂法对 A 接近于 $\alpha=-1.4$ 的特征值进行估值，计算精确到 4 位小数.

[M]在习题 19 和习题 20 中，求（a）最大特征值，（b）接近于零的特征值. 设 $\boldsymbol{x}_0=(1,0,0,0)$，在求特征值时，要求近似序列的计算精确到 4 位小数，还包括近似特征向量的计算.

19. $A=\begin{bmatrix}10&7&8&7\\7&5&6&5\\8&6&10&9\\7&5&9&10\end{bmatrix}$
20. $A=\begin{bmatrix}1&2&3&2\\2&12&13&11\\-2&3&0&2\\4&5&7&2\end{bmatrix}$
21. 一个常见的误解是：假如 A 有严格主特征值，那么对足够大的 k 值，向量 $A^k\boldsymbol{x}$ 近似于 A 的某个特征向量. 对下面给出的三个矩阵，当 $\boldsymbol{x}=(0.5,0.5)$ 时，计算并研究 $A^k\boldsymbol{x}$，并试图得到一般结论（对 2×2 矩阵）.

    a. $A=\begin{bmatrix}0.8&0\\0&0.2\end{bmatrix}$

    b. $A=\begin{bmatrix}1&0\\0&0.8\end{bmatrix}$

    c. $A=\begin{bmatrix}8&0\\0&2\end{bmatrix}$

练习题答案

对给出的 A 和 $\boldsymbol{x}$，

$$A\boldsymbol{x}=\begin{bmatrix}5&8&4\\8&3&-1\\4&-1&2\end{bmatrix}\begin{bmatrix}1.00\\-4.30\\8.10\end{bmatrix}=\begin{bmatrix}3.00\\-13.00\\24.50\end{bmatrix}$$

如果 $A\boldsymbol{x}$ 近似于 $\boldsymbol{x}$ 的数倍，那么两个向量相应分量的比率也应该近似于某个常量，故计算：

{$A\boldsymbol{x}$ 的元素}	÷{$\boldsymbol{x}$ 的元素}	={比例}
3.00	1.00	3.000
−13.00	−4.30	3.023
24.50	8.10	3.025

$A\boldsymbol{x}$ 的每一个分量大约是 $\boldsymbol{x}$ 对应分量的 3 倍，故 $\boldsymbol{x}$ 接近于 A 的特征向量，上面比例的任何一个都是特征值的近似值.（若精确到 5 位小数，特征值是 3.024 09.）

补充习题

在下列所有补充习题中，A 和 B 均为适当大小的方阵.

1. 判断命题的真假，给出理由.

    a. 设 A 可逆且 1 是其特征值，则 1 也是 A^{-1} 的

特征值.

b. 若 A 可通过行变换化成单位阵 I，那么 A 可对角化.

c. 若 A 的某一行或某一列上元素全为 0，那么 0 是 A 的一个特征值.

d. A 的每个特征值同样也是 A^2 的特征值.

e. A 的每个特征向量同样也是 A^2 的特征向量.

f. 可逆矩阵 A 的每个特征向量同时也是 A^{-1} 的特征向量.

g. 特征值必为非零标量.

h. 特征向量必为非零向量.

i. 对应同一特征值的两个特征向量必为线性相关的.

j. 相似矩阵必有相同的特征值.

k. 相似矩阵必有相同的特征向量.

l. 矩阵 A 的两个特征向量之和仍是 A 的特征向量.

m. 上三角矩阵 A 的特征值正好是对角矩阵 A 中的非零元素.

n. 将重数计算在内，矩阵 A 和 A^T 具有相同的特征值.

o. 如果 5×5 矩阵 A 的不同特征值少于 5 个，那么 A 不可对角化.

p. $\mathbb{R}^2$ 中存在没有特征向量的 2×2 矩阵.

q. 若 A 可对角化，那么 A 的各列线性无关.

r. 一个非零向量不能对应于 A 的两个不同特征值.

s. 方阵 A 可逆当且仅当存在一个坐标系使得变换 $\mathbf{x} \mapsto A\mathbf{x}$ 可以用对角矩阵来表示.

t. 如果 $\mathbb{R}^n$ 的标准基中每一向量 $\mathbf{e}_j$ 都是 A 的特征向量，那么 A 是一对角矩阵.

u. 如果 A 相似于一可对角化的矩阵 B，那么 A 也可对角化.

v. 如果 A 和 B 均为 $n \times n$ 的可逆矩阵，那么 AB 相似于 BA.

w. 具有 n 个线性无关特征向量的 $n \times n$ 矩阵可逆.

x. 如果 A 是一 $n \times n$ 可对角化矩阵，那么 $\mathbb{R}^n$ 中的任一向量均可表为 A 的特征向量的线性组合.

2. $\mathbf{x}$ 是矩阵乘积 AB 的一个特征向量，且 $B\mathbf{x} \neq \mathbf{0}$，证明 $B\mathbf{x}$ 是 BA 的一个特征向量.

3. 设 $\mathbf{x}$ 是 A 的对应于特征值 λ 的特征向量.

   a. 证明 $\mathbf{x}$ 是 $5I - A$ 的一个特征向量. 并求其对应的特征值.

   b. 证明 $\mathbf{x}$ 是 $5I - 3A + A^2$ 的一个特征向量. 并求其对应的特征值.

4. 用数学归纳法证明如果 λ 是 $n \times n$ 矩阵 A 的一个特征值，$\mathbf{x}$ 是其对应的特征向量，那么对任一正整数 m，λ^m 是 A^m 的一个特征值，$\mathbf{x}$ 是其对应的特征向量.

5. 设 $p(t) = c_0 + c_1 t + c_2 t^2 + \cdots + c_n t^n$，定义 $p(A)$ 为用对应的 A 的幂去取代 $p(t)$ 中的每一个 t 的幂所得的矩阵（其中 $A^0 = I$）. 即

$$p(A) = c_0 I + c_1 A + c_2 A^2 + \cdots + c_n A^n$$

   若 λ 是 A 的一个特征值，证明 $p(A)$ 的一个特征值是 $p(\lambda)$.

6. 设 $A = PDP^{-1}$, P 是 2×2 矩阵，$D = \begin{bmatrix} 2 & 0 \\ 0 & 7 \end{bmatrix}$.

   a. 设 $B = 5I - 3A + A^2$. 证明 B 可通过寻找 B 的一个合适的因式分解来化成对角矩阵.

   b. 已知 $p(t)$ 和 $p(A)$ 如习题 5 所述，证明 $p(A)$ 可对角化.

7. 设 A 可对角化且 $p(t)$ 是 A 的特征多项式. 定义 $p(A)$ 如习题 5 所述，证明 $p(A)$ 是零矩阵. 这一事实，叫做 Cayley-Hamilton 定理，它对任一方阵均成立.

8. a. 设 A 是一 $n \times n$ 的可对角化矩阵. 证明当特征值 λ 的重数是 n 时，有 $A = \lambda I$.

   b. 利用（a）来证明矩阵 $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$ 不可对角化.

9. 证明：当 A 的所有特征值均小于 1 时，$I - A$ 可逆.（提示：如果 $I - A$ 不可逆情况会是怎样?）

10. 证明：若 A 可对角化，且 A 的所有特征值均小于 1，那么当 $k\to\infty$ 时，A^k 趋于零矩阵.（提示：考虑 $A^k\boldsymbol{x}$，其中 $\boldsymbol{x}$ 为 I 的任一列.）
11. 设 $\boldsymbol{u}$ 为 A 的对应于特征值 λ 的特征向量，H 为 $\mathbb{R}^n$ 上过 $\boldsymbol{u}$ 和原点的直线.
    a. 证明：对 H 上任一 $\boldsymbol{x}$，有 $A\boldsymbol{x}$ 在 H 上，那么 H 在 A 内不变.
    b. 设 K 为 $\mathbb{R}^n$ 上恒定于 A 内的一维子空间，说明：K 包含 A 的一个特征向量.
12. 设 $G=\begin{bmatrix}A & X\\ 0 & B\end{bmatrix}$.利用 5.2 节中求行列式的公式（1）证明 $\det G=(\det A)(\det B)$. 由此，推导出 G 的特征多项式为 A 与 B 的特征多项式的乘积.

利用习题 12 求下列矩阵的特征值.

13. $A=\begin{bmatrix}3 & -2 & 8\\ 0 & 5 & -2\\ 0 & -4 & 3\end{bmatrix}$
14. $A=\begin{bmatrix}1 & 5 & -6 & -7\\ 2 & 4 & 5 & 2\\ 0 & 0 & -7 & -4\\ 0 & 0 & 3 & 1\end{bmatrix}$
15. 设 J 为元素值全为 1 的 $n\times n$ 矩阵，并设 $A=(a-b)I+bJ$，即 $A=\begin{bmatrix}a & b & b & \cdots & b\\ b & a & b & \cdots & b\\ b & b & a & \cdots & b\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ b & b & b & \cdots & a\end{bmatrix}$，利用第 3 章补充习题 16 的结果证明 A 的特征值为 $a-b$ 和 $a+(n-1)b$. 并求这两个特征值的重数.
16. 应用习题 15 的结论求矩阵 $\begin{bmatrix}1 & 2 & 2\\ 2 & 1 & 2\\ 2 & 2 & 1\end{bmatrix}$ 和 $\begin{bmatrix}7 & 3 & 3 & 3 & 3\\ 3 & 7 & 3 & 3 & 3\\ 3 & 3 & 7 & 3 & 3\\ 3 & 3 & 3 & 7 & 3\\ 3 & 3 & 3 & 3 & 7\end{bmatrix}$ 的特征值.
17. 设 $A=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}$. 回顾 5.4 节习题 25 中的 $\operatorname{tr} A$（A 的轨迹）等于 A 的对角元素之和. 证明：A 的特征多项式为 $\lambda^2-(\operatorname{tr} A)\lambda+\det A$. 然后证明一 2×2 矩阵 A 的特征值全为实数当且仅当 $\det A\leqslant\left(\dfrac{\operatorname{tr} A}{2}\right)^2$.
18. 设 $A=\begin{bmatrix}0.4 & -0.3\\ 0.4 & 1.2\end{bmatrix}$. 说明当 $k\to\infty$ 时，$A^k\to\begin{bmatrix}-0.5 & -0.75\\ 1.0 & 1.50\end{bmatrix}$.

习题 19~23 是关于特征多项式 $p(t)=a_0+a_1t+\cdots+a_{n-1}t^{n-1}+t^n$ 和 $n\times n$ 矩阵 C_p 的，C_p 叫做 p 的**友矩阵**：

$$C_p=\begin{bmatrix}0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & & 0\\ \vdots & & & & \vdots\\ 0 & 0 & 0 & & 1\\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1}\end{bmatrix}.$$

19. $p(t)=6-5t+t^2$，写出它的友矩阵 C_p，并求 C_p 的特征多项式.
20. 设 $p(t)=(t-2)(t-3)(t-4)=-24+26t-9t^2+t^3$，写出 $p(t)$ 的友矩阵，并用第 3 章的方法求它的特征多项式.
21. 利用数学归纳法证明：当 $n\geqslant 2$ 时，
$$\det(C_p-\lambda I)=(-1)^n(a_0+a_1\lambda+\cdots+a_{n-1}\lambda^{n-1}+\lambda^n)=(-1)^n p(\lambda)$$
（提示：求沿第一列展开的余子式，证明 $\det(C_p-\lambda I)$ 具有 $(-\lambda)B+(-1)^n a_0$ 的形式，其中 B 为某一多项式（由归纳假设）.）
22. 设 $p(t)=a_0+a_1t+a_2t^2+t^3$，并设 λ 是 p 的一个零点.
    a. 写出 p 的友矩阵.
    b. 证明 $\lambda^3=-a_0-a_1\lambda-a_2\lambda^2$ 及 $(1,\lambda,\lambda^2)$ 是 p 的友矩阵的一个特征向量.
23. 设 p 为习题 22 中的多项式，并假设方程 $p(t)=0$ 有不同的根 $\lambda_1,\lambda_2,\lambda_3$. 设 V 为范德蒙德矩阵 $V=\begin{bmatrix}1 & 1 & 1\\ \lambda_1 & \lambda_2 & \lambda_3\\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2\end{bmatrix}$.（$V$ 的转置在第 2 章

补充习题 11 提及.）利用习题 22 和本章中的定理推导 V 是可逆的（不要计算 V^{-1}）. 然后说明 $V^{-1}C_pV$ 是一个对角矩阵.

24. MATLAB 命令 roots(p)是用来计算多项式方程 $p(t)=0$ 的根的. 阅读 MATLAB 手册，阐述 roots 命令所用到的算法的基本思想.

25. [M]如果可能的话，利用矩阵程序对角化矩阵 $A=\begin{bmatrix} -3 & -2 & 0 \\ 14 & 7 & -1 \\ -6 & -3 & 1 \end{bmatrix}$. 利用特征值命令创建对角矩阵 D. 如果程序有产生特征向量的命令，利用此命令产生的向量来构造可逆矩阵 P. 然后计算 $AP-PD$ 和 PDP^{-1}. 并讨论结果.

26. [M]设 $A=\begin{bmatrix} -8 & 5 & -2 & 0 \\ -5 & 2 & 1 & -2 \\ 10 & -8 & 6 & -3 \\ 3 & -2 & 1 & 0 \end{bmatrix}$，重做习题 25.

第6章 正交性和最小二乘法

介绍性实例 北美数据 GPS 导航

设想启动一个巨大的工程，该工程估计持续 10 年时间，需要花费大量的人力去构造并计算一个 1 800 000×900 000 的线性方程组. 这实际上就是美国地质调查局（NGS）1974 年所做的工作，当时准备更新北美地质资料（NAD）——一个包含 268 000 个仔细测量并标记说明的地点网络，它覆盖整个北美大陆，包括巴拿马地峡、格陵兰岛、夏威夷、维尔京群岛、波多黎各和其他加勒比海诸岛.

北美地质资料中记录的经度和纬度范围必须精确到几厘米，其原因是它构成诸如测量、地图、法定边界、国家和区域土地使用计划，像高速公路和公共使用线路等国内工程项目的设计等方面的标准. 然而从上次 1927 年地名测量调整以来，至少需要在原始资料中增加 200 000 个新观测点，由于不严密的测量和地球板块的飘移，误差已逐年积累. 至于北美地质资料重新调整所需要的数据采集工作，1983 年就已完成.

通常意义下，北美地质资料对应的线性方程组没有正常解，只具有最小二乘解. 它用经度和纬度表示参考点以逼近 180 万个实际观测点. 这种最小二乘解发现于 1986 年，是通过求解一个相关的所谓法方程组而得，其计算涉及 928 735 个方程、928 735 个变量！[1]

最近，为了能通过基于人造地球卫星的全球定位系统（GPS）给人造地球卫星以精确定位，地面上的参考点的知识已变得至关重要. 一个配有 GPS 装置的卫星能通过测量从地面上三个发射器发射的信号的到达时间来计算它相对于地球的位置. 为此，卫星使用了与地面站同步精确的原子钟， 而这些地面站的位置则是通过 NAD 精确获知了的.

全球定位系统既可用于定位地面上的新参考点，也可用于发现一个使用者相对于给定地图的位置. 当汽车司机（或登山员）打开一个 GPS 接收器时，这个接收器便会测量至少从三个卫星发来的信号的相对到达时间. 这些信息，连同发过来的关于卫星的位置及相应时刻，将用于调整 GPS 接收器的

[1] 有关 Helmert 分块技巧的数学讨论和整个 NAD 项目的详细内容，出现在 *North American Datum of 1983*, Charles R.Schwarz(ed.), National Geodetic Survey, National Oceanic and Atmospheric Administration (NOAA) Professional Paper NOS 2，1989.

时间并确定它在地球上的近似位置. 如果有从第 4 颗卫星发来的信息，GPS 接收器甚至还能确定它自己的近似高度.

不管是 NAD 问题还是 GPS 问题，都是通过寻找一个“近似满足”不相容线性方程组的向量得以解决. 这种似乎矛盾的说法将需要本章前 5 节建立起来的思想得以详细解释. >>>>>>>>

为了寻找一个没有真正解的不相容的线性方程组的近似解，需要给“接近”概念下一个良好的定义. 6.1 节介绍向量空间中的距离和正交性概念. 6.2 节和 6.3 节说明正交性是如何用于判定子空间 W 中的某个点是最接近 W 之外的一个给定点 y 的. 通过把 W 设定为矩阵的列子空间，6.5 节导出了一个可以求得不相容线性方程组（NAD 报告中所解答的正是这样的方程组）的近似（“最小二乘”）解的方法.

6.4 节让我们又一次看到正交投影的作用，导出了一个广泛应用于数值线性代数中的矩阵分解方法. 本章其余节的内容是验算实际应用中产生的一部分最小二乘问题，有些涉及比 $\mathbb{R}^n$ 空间更一般的向量空间.

6.1 内积、长度和正交性

大家已经熟悉了二维和三维空间中的长度、距离和垂直等几何概念，本节类似引入 $\mathbb{R}^n$ 空间的定义，这些概念为解决实际问题（如上面提到的最小二乘问题）提供了有力的几何工具. 而 $\mathbb{R}^n$ 中的三个新概念都建立在两个向量的内积基础之上.

内积

如果 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是 $\mathbb{R}^n$ 空间中的向量，可以将 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 作为 $n\times 1$ 矩阵. 向量矩阵 $\boldsymbol{u}^{\mathrm{T}}$ 是 $1\times n$ 矩阵且矩阵乘积 $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}$ 是一个 1×1 矩阵，我们将其记为一个不加括号的实数. 如 $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}$ 称为 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的内积，并记作 $\boldsymbol{u}\cdot\boldsymbol{v}$. 这里的内积，曾在练习 2.1 提到过，也称为**点积**. 如果

$$\boldsymbol{u}=\begin{bmatrix}u_1\\u_2\\\vdots\\u_n\end{bmatrix},\quad \boldsymbol{v}=\begin{bmatrix}v_1\\v_2\\\vdots\\v_n\end{bmatrix}$$

那么 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的内积定义为：

$$\begin{bmatrix}u_1 & u_2 & \cdots & u_n\end{bmatrix}\begin{bmatrix}v_1\\v_2\\\vdots\\v_n\end{bmatrix}=u_1v_1+u_2v_2+\cdots+u_nv_n$$

例 1 如果 $\boldsymbol{u}=\begin{bmatrix}2\\-5\\-1\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}3\\2\\-3\end{bmatrix}$，计算 $\boldsymbol{u}\cdot\boldsymbol{v}$ 和 $\boldsymbol{v}\cdot\boldsymbol{u}$.

解

$$u\cdot v=u^{T}v=[2\quad -5\quad -1]\begin{bmatrix}3\\2\\-3\end{bmatrix}=(2)(3)+(-5)(2)+(-1)(-3)=-1$$

$$v\cdot u=v^{T}u=[3\quad 2\quad -3]\begin{bmatrix}2\\-5\\-1\end{bmatrix}=(3)(2)+(2)(-5)+(-3)(-1)=-1$$

■

从例 1 中的计算明显看出，$u\cdot v=v\cdot u$. 一般情形下，内积成立交换律. 下面关于内积的性质，可以很容易用 2.1 节中的矩阵转置运算来验证（见本节末习题 21 和 22）.

定理 1 设 v,u 和 w 是 $\mathbb{R}^n$ 空间中的向量，c 是一个数，那么

a. $u\cdot v= v\cdot u$.

b. $(u+v)\cdot w=u\cdot w+v\cdot w$.

c. $(cu)\cdot v=c(u\cdot v)=u\cdot(cv)$.

d. $u\cdot u\geqslant 0$, 并且 $u\cdot u=0$ 成立的充分必要条件是 $u=0$.

性质（b）和（c）可以合并为以下法则：

$$(c_1u_1+\cdots+c_pu_p)\cdot w=c_1(u_1\cdot w)+\cdots+c_p(u_p\cdot w)$$

向量的长度

如果 v 是属于 $\mathbb{R}^n$ 的向量，其分量为 $v_1,\cdots,v_n$，因为 $v\cdot v$ 是非负数，那么 $v\cdot v$ 的平方根有意义.

定义 向量 v 的长度（范数）是非负数 $\|v\|$，定义为

$$\|v\|=\sqrt{v\cdot v}=\sqrt{v_1^2+v_2^2+\cdots+v_n^2}\ \text{ 且 }\|v\|^2=v\cdot v$$

假若 v 是 $\mathbb{R}^2$ 中的向量，且 $v=\begin{bmatrix}a\\b\end{bmatrix}$. 如果我们将 v 与平面上的点 (a, b) 相对应，那么 $\|v\|$ 的值和平面内原点到点 v 的线段长度一致，这个结论可以从勾股定理和图 6-1 中的三角形直接得到.

类似长方体的对角线计算，三维空间中向量 $\|v\|$ 的长度和通常意义下的长度概念是一致的.

对任意数 c，向量 cv 的长度等于 $|c|$ 乘 v 的长度，即

$$\|cv\|=|c|\ \|v\|$$

（为验证上式，计算 $\|cv\|^2=(cv)\cdot(cv)=c^2v\cdot v=c^2\|v\|^2$，然后作开方运算就可以得到以上结论.）

图 6-1 $\|v\|$ 作为长度的几何意义

长度为 1 的向量称为**单位向量**，如果把一个非零向量除以自身的长度，即乘 $\frac{1}{\|v\|}$，就可以得到一个单位化的向量，$u=\frac{v}{\|v\|}$，因为 u 的长度是 $\left(1/\|v\|\right)\|v\|$. 这种把向量 v 化成单位向量 u 的过程，称为向量 v 的**单位化**，且此时 u 和 v 方向一致.

下面用简约形式（即行向量形式）计算几个例题.

例2　若 $\boldsymbol{v}=(1,-2,2,0)$，找出和 $\boldsymbol{v}$ 方向一致的单位向量 $\boldsymbol{u}$.

解　首先计算向量 $\boldsymbol{v}$ 的长度

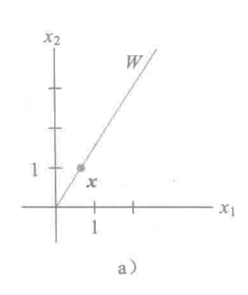
$$\|\boldsymbol{v}\|^2=\boldsymbol{v}\cdot\boldsymbol{v}=(1)^2+(-2)^2+(2)^2+(0)^2=9,\quad \|\boldsymbol{v}\|=\sqrt{9}=3$$

对 $\boldsymbol{v}$ 乘 $\frac{1}{\|\boldsymbol{v}\|}$ 得到

$$\boldsymbol{u}=\frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}=\frac{1}{3}\boldsymbol{v}=\frac{1}{3}\begin{bmatrix}1\\-2\\2\\0\end{bmatrix}=\begin{bmatrix}1/3\\-2/3\\2/3\\0\end{bmatrix}$$

为验证 $\|\boldsymbol{u}\|=1$，只需验证 $\|\boldsymbol{u}\|^2=1$.

$$\|\boldsymbol{u}\|^2=\boldsymbol{u}\cdot\boldsymbol{u}=(\tfrac{1}{3})^2+(-\tfrac{2}{3})^2+(\tfrac{2}{3})^2+(0)^2=\frac{1}{9}+\frac{4}{9}+\frac{4}{9}+0=1$$ ■



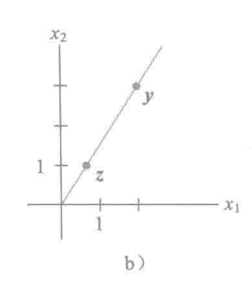


图 6-2　把一个向量单位化，得到一个单位向量

例3　设 W 是 $\mathbb{R}^n$ 的子空间且由向量 $\boldsymbol{x}=(\frac{2}{3},1)$ 生成，求出一个单位向量 $\boldsymbol{z}$ 且 $\boldsymbol{z}$ 构成 W 的一个基.

解　空间 W 包含所有 $\boldsymbol{x}$ 数倍的向量，如图 6-2a 所示. W 中的任意非零向量都是 W 的基. 为简化计算，重新"标度" $\boldsymbol{x}$ 以消去分数，即向量 $\boldsymbol{x}$ 乘 3 得到 $\boldsymbol{y}=\begin{bmatrix}2\\3\end{bmatrix}$，现在计算 $\|\boldsymbol{y}\|^2=2^2+3^2=13$，$\|\boldsymbol{y}\|=\sqrt{13}$. 把向量 $\boldsymbol{y}$ 单位化可得：

$$\boldsymbol{z}=\frac{1}{\sqrt{13}}\begin{bmatrix}2\\3\end{bmatrix}=\begin{bmatrix}\frac{2}{\sqrt{13}}\\ \frac{3}{\sqrt{13}}\end{bmatrix}$$

见图 6-2b. 另外一个单位向量是 $(-2/\sqrt{13},\,-3/\sqrt{13})$. ■

$\mathbb{R}^n$ 空间中的距离

接下来描述一个向量如何逼近另一个向量，注意，如果 a 和 b 是实数，在数轴上 a 与 b 的距离是 $|a-b|$，图 6-3 给出两个实例. 用类似 $\mathbb{R}$ 中两个数间的距离定义 $\mathbb{R}^n$ 空间中两个向量的距离.

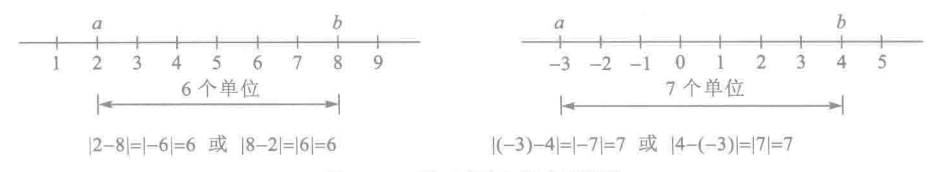


图 6-3　$\mathbb{R}$ 中两个数的距离

定义　$\mathbb{R}^n$ **中向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的距离，记作 $\text{dist}(\boldsymbol{u},\boldsymbol{v})$，表示向量 $\boldsymbol{u}-\boldsymbol{v}$ 的长度，即**

$$\text{dist}(\boldsymbol{u},\boldsymbol{v})=\|\boldsymbol{u}-\boldsymbol{v}\|$$

对 $\mathbb{R}^2$ 和 $\mathbb{R}^3$ 空间，如例 4、例 5 所示，距离的定义和欧几里得空间中两点的距离公式一致.

例 4 计算向量 $\boldsymbol{u}=(7, 1)$ 和 $\boldsymbol{v}=(3, 2)$ 之间的距离.

解 先计算 $\boldsymbol{u}-\boldsymbol{v}=\begin{bmatrix}7\\1\end{bmatrix}-\begin{bmatrix}3\\2\end{bmatrix}=\begin{bmatrix}4\\-1\end{bmatrix}$，则有 $\|\boldsymbol{u}-\boldsymbol{v}\|=\sqrt{4^2+(-1)^2}=\sqrt{17}$.

向量 $\boldsymbol{u}$，$\boldsymbol{v}$ 和 $\boldsymbol{u}-\boldsymbol{v}$ 如图 6-4 所示，向量 $\boldsymbol{u}-\boldsymbol{v}$ 加上向量 $\boldsymbol{v}$ 的结果是向量 $\boldsymbol{u}$. 注意到图 6-4 中的平行四边形表明，从向量 $\boldsymbol{u}$ 到 $\boldsymbol{v}$ 的距离与向量 $\boldsymbol{u}-\boldsymbol{v}$ 到 $\boldsymbol{0}$ 的距离相等.

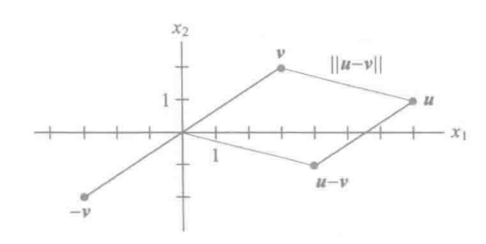


图 6-4 向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的距离等于 $\boldsymbol{u}-\boldsymbol{v}$ 的长度 ■

例 5 如果 $\boldsymbol{u}=(u_1, u_2, u_3)$ 和 $\boldsymbol{v}=(v_1, v_2, v_3)$，那么

$$\operatorname{dist}(\boldsymbol{u}, \boldsymbol{v})=\|\boldsymbol{u}-\boldsymbol{v}\|=\sqrt{(\boldsymbol{u}-\boldsymbol{v})\cdot(\boldsymbol{u}-\boldsymbol{v})}=\sqrt{(u_1-v_1)^2+(u_2-v_2)^2+(u_3-v_3)^2}$$ ■

正交向量

本章以下内容阐述这样的事实，即欧几里得空间中的直线垂直概念可以推广到 $\mathbb{R}^n$ 空间.

考虑 $\mathbb{R}^2$ 或 $\mathbb{R}^3$ 空间中，通过原点且由向量 $\boldsymbol{u}$ 和向量 $\boldsymbol{v}$ 确定的两条直线，两条直线（见图 6-5）几何上垂直当且仅当从 $\boldsymbol{u}$ 到 $\boldsymbol{v}$ 的距离与从 $\boldsymbol{u}$ 到 $-\boldsymbol{v}$ 的距离相等，这等同于要求它们距离的平方要相等.

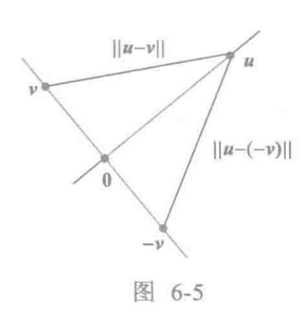


图 6-5

现在计算：

$$\begin{aligned}
[\operatorname{dist}(\boldsymbol{u}, -\boldsymbol{v})]^2 &= \|\boldsymbol{u}-(-\boldsymbol{v})\|^2=\|\boldsymbol{u}+\boldsymbol{v}\|^2 \\
&= (\boldsymbol{u}+\boldsymbol{v})\cdot(\boldsymbol{u}+\boldsymbol{v}) \\
&= \boldsymbol{u}\cdot(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{v}\cdot(\boldsymbol{u}+\boldsymbol{v}) && \text{定理 1 (b)} \\
&= \boldsymbol{u}\cdot\boldsymbol{u}+\boldsymbol{u}\cdot\boldsymbol{v}+\boldsymbol{v}\cdot\boldsymbol{u}+\boldsymbol{v}\cdot\boldsymbol{v} && \text{定理 1 (a)、(b)}
\end{aligned}$$

$$=\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2+2\boldsymbol{u}\cdot\boldsymbol{v} \qquad \text{定理 1（a）} \qquad (1)$$

同样将 $-\boldsymbol{v}$ 换成 $\boldsymbol{v}$ 的计算如下：

$$\begin{aligned}[\operatorname{dist}(\boldsymbol{u},\boldsymbol{v})]^2&=\|\boldsymbol{u}\|^2+\|-\boldsymbol{v}\|^2+2\boldsymbol{u}\cdot(-\boldsymbol{v})\\&=\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2-2\boldsymbol{u}\cdot\boldsymbol{v}\end{aligned}$$

两个距离平方相等的充分必要条件是：$2\boldsymbol{u}\cdot\boldsymbol{v}=-2\boldsymbol{u}\cdot\boldsymbol{v}$ 或 $\boldsymbol{u}\cdot\boldsymbol{v}=0$．

这里的计算表明，当向量 $\boldsymbol{u}$ 和向量 $\boldsymbol{v}$ 看作几何点时，通过这些点和原点的两条直线相互垂直的充分必要条件是 $\boldsymbol{u}\cdot\boldsymbol{v}=0$．下面给出 $\mathbb{R}^n$ 空间中两个向量互相垂直的一般定义（或正交，这是线性代数中的一个通用术语）．

定义 如果 $\boldsymbol{u}\cdot\boldsymbol{v}=0$，则 $\mathbb{R}^n$ 中的两个向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 称为（相互）**正交的**.

由 $\boldsymbol{0}^{\mathrm{T}}\cdot\boldsymbol{v}=0$ 对任意 $\boldsymbol{v}$ 都成立，可以得出零向量与任意向量正交.

关于向量正交的一个重要性质由下面的定理给出，其证明可以从上面正交性定义和（1）中的计算立刻推出. 图 6-6 的直角三角形给出长度的直观描述.

定理 2 （毕达哥拉斯（勾股）定理）

两个向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 正交的充分必要条件是 $\|\boldsymbol{u}+\boldsymbol{v}\|^2=\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2$.

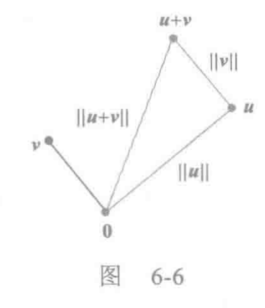


图 6-6

正交补

为了提供内积概念的相关练习，我们引入一个在 6.3 节和其余章节都需要的概念. 如果向量 $\boldsymbol{z}$ 与 $\mathbb{R}^n$ 的子空间 W 中的任意向量都正交，则称 $\boldsymbol{z}$ **正交于** W．与子空间 W 正交的向量 $\boldsymbol{z}$ 的全体组成的集合称为 W 的**正交补**，并记作 $W^{\perp}$（$W^{\perp}$ 读作 W 正交补）.

例 6 设 W 是 $\mathbb{R}^3$ 空间中通过原点 $\boldsymbol{O}$ 的平面，L 是通过原点且与 W 正交的直线. 如果 $\boldsymbol{z}$ 和 $\boldsymbol{w}$ 都非零，$\boldsymbol{z}$ 在直线 L 上且 $\boldsymbol{w}$ 在空间 W 内，那么从 $\boldsymbol{0}$ 到 $\boldsymbol{z}$ 的线段正交于从 $\boldsymbol{0}$ 到 $\boldsymbol{w}$ 的线段，即 $\boldsymbol{z}\cdot\boldsymbol{w}=0$，见图 6-7. 从而 L 上的每个向量与 W 中的任一向量都正交. 事实上，L 包含所有与空间 W 中的向量 $\boldsymbol{w}$ 都正交的全体向量，空间 W 包含与 L 中向量 $\boldsymbol{z}$ 都正交的全体向量，也就是说：

$$L=W^{\perp} \text{ 且 } W=L^{\perp}$$ ■

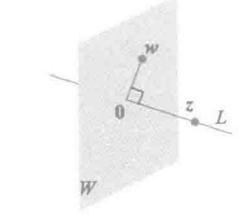


图 6-7 一个作为正交补空间的平面和通过原点的直线

若 W 是 $\mathbb{R}^n$ 的子空间，下面两个关于 $W^{\perp}$ 的性质，会在以后的章节用到，证明放在习题 29 和习题 30 中，习题 27~31 将给出几个复习内积性质的练习.

1. 向量 $\boldsymbol{x}$ 属于 $W^{\perp}$ 的充分必要条件是向量 $\boldsymbol{x}$ 与生成空间 W 的任一向量都正交.
2. $W^{\perp}$ 是 $\mathbb{R}^n$ 的一个子空间.

下面的定理和习题 31 验证了 4.6 节关于子空间的论断，见图 6-8（参考 4.6 节的习题 28）.

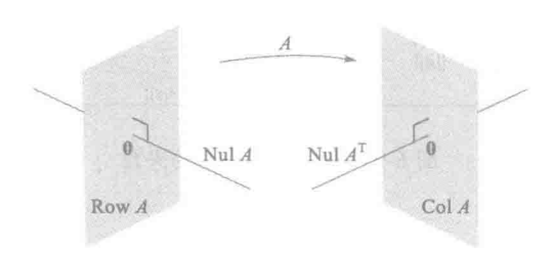


图 6-8 一个由 $m\times n$ 矩阵 A 确定的基本子空间

定理 3 假设 A 是 $m\times n$ 矩阵，那么 A 的行向量空间的正交补空间是 A 的零空间，且 A 的列向量空间的正交补是 A^T 的零空间：

$$(\text{Row } A)^{\perp} = \text{Nul } A \quad \text{且} \quad (\text{Col } A)^{\perp} = \text{Nul } A^T$$

证 计算 $A\boldsymbol{x}$ 的行列法则表明，如果 $\boldsymbol{x}$ 是 Nul A 中的向量，那么向量 $\boldsymbol{x}$ 与 A 的每一行（将行作为 $\mathbb{R}^n$ 空间中的向量）正交. 由于 A 的行生成行空间，向量 $\boldsymbol{x}$ 与 Row A 正交. 反之，如果 $\boldsymbol{x}$ 与 Row A 正交，那么 $\boldsymbol{x}$ 当然与 A 的每一行正交，因此 $A\boldsymbol{x}=\boldsymbol{0}$，从而证明了第一个结论. 如果将 A 换成 A^T 并利用 Row A^T = Col A，可以用第一个结论的证明得出第二个结论. ■

$\mathbb{R}^2$ 空间和 $\mathbb{R}^3$ 空间的角度（可选内容）

如果 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是 $\mathbb{R}^2$ 或 $\mathbb{R}^3$ 中的非零向量，那么可以用内积，将从原点到点 $\boldsymbol{u}$ 和原点到点 $\boldsymbol{v}$ 的两个线段之间的夹角联系起来，对应的公式是：

$$\boldsymbol{u}\cdot\boldsymbol{v} = \|\boldsymbol{u}\|\ \|\boldsymbol{v}\|\cos\vartheta \tag{2}$$

为了验证 $\mathbb{R}^2$ 中的向量公式，考虑图 6-9 所示的三角形，其边长分别是 $\|\boldsymbol{u}\|$, $\|\boldsymbol{v}\|$ 和 $\|\boldsymbol{u}-\boldsymbol{v}\|$，由余弦定理可知：

$$\|\boldsymbol{u}-\boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 - 2\|\boldsymbol{u}\|\ \|\boldsymbol{v}\|\cos\vartheta$$

可以重新组合上面的内积表达式：

$$\begin{aligned}\|\boldsymbol{u}\|\ \|\boldsymbol{v}\|\cos\vartheta &= \frac{1}{2}\left[\|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2\right] \\ &= \frac{1}{2}[u_1^2 + u_2^2 + v_1^2 + v_2^2 - (u_1-v_1)^2 - (u_2-v_2)^2] \\ &= u_1v_1 + u_2v_2 \\ &= \boldsymbol{u}\cdot\boldsymbol{v}\end{aligned}$$

$\mathbb{R}^3$ 空间的情形可类似验证. 当 $n>3$ 时，公式（2）可用于定义两个向量之间的夹角. 例如，在统计学中，（2）式中对向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 定义的 $\cos\vartheta$ 的值就是统计学家所称的相关系数.

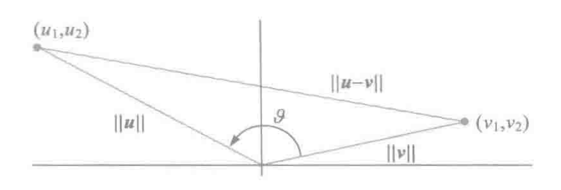


图 6-9 两个向量之间的夹角

练习题

令 $\boldsymbol{a}=\begin{bmatrix}-2\\1\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}-3\\1\end{bmatrix}$, $\boldsymbol{c}=\begin{bmatrix}\frac{4}{3}\\-1\\\frac{2}{3}\end{bmatrix}$, $\boldsymbol{d}=\begin{bmatrix}5\\6\\-1\end{bmatrix}$.

1. 计算 $\dfrac{\boldsymbol{a}\cdot\boldsymbol{b}}{\boldsymbol{a}\cdot\boldsymbol{a}}$ 和 $\left(\dfrac{\boldsymbol{a}\cdot\boldsymbol{b}}{\boldsymbol{a}\cdot\boldsymbol{a}}\right)\boldsymbol{a}$.
2. 计算向量 $\boldsymbol{c}$ 方向的单位向量 $\boldsymbol{u}$.
3. 证明向量 $\boldsymbol{d}$ 和向量 $\boldsymbol{c}$ 正交.
4. 利用练习题 2 和 3 的结果，解释为什么 $\boldsymbol{d}$ 一定正交于单位向量 $\boldsymbol{u}$.

习题 6.1

在习题 1~8 中，利用下列向量计算数值：

$$\boldsymbol{u}=\begin{bmatrix}-1\\2\end{bmatrix},\quad \boldsymbol{v}=\begin{bmatrix}4\\6\end{bmatrix},\quad \boldsymbol{w}=\begin{bmatrix}3\\-1\\-5\end{bmatrix},\quad \boldsymbol{x}=\begin{bmatrix}6\\-2\\3\end{bmatrix}$$

1. $\boldsymbol{u}\cdot\boldsymbol{u}$, $\boldsymbol{v}\cdot\boldsymbol{u}$ 和 $\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{\boldsymbol{u}\cdot\boldsymbol{u}}$
2. $\boldsymbol{w}\cdot\boldsymbol{w}$, $\boldsymbol{x}\cdot\boldsymbol{w}$ 和 $\dfrac{\boldsymbol{x}\cdot\boldsymbol{w}}{\boldsymbol{w}\cdot\boldsymbol{w}}$
3. $\dfrac{1}{\boldsymbol{w}\cdot\boldsymbol{w}}\boldsymbol{w}$
4. $\dfrac{1}{\boldsymbol{u}\cdot\boldsymbol{u}}\boldsymbol{u}$
5. $\left(\dfrac{\boldsymbol{u}\cdot\boldsymbol{v}}{\boldsymbol{v}\cdot\boldsymbol{v}}\right)\boldsymbol{v}$
6. $\left(\dfrac{\boldsymbol{x}\cdot\boldsymbol{w}}{\boldsymbol{x}\cdot\boldsymbol{x}}\right)\boldsymbol{x}$
7. $\|\boldsymbol{w}\|$
8. $\|\boldsymbol{x}\|$

在习题 9~12 中，计算给定向量方向的单位向量.

9. $\begin{bmatrix}-30\\40\end{bmatrix}$
10. $\begin{bmatrix}-6\\4\\-3\end{bmatrix}$
11. $\begin{bmatrix}7/4\\1/2\\1\end{bmatrix}$
12. $\begin{bmatrix}8/3\\2\end{bmatrix}$

13. 计算向量 $\boldsymbol{x}=\begin{bmatrix}10\\-3\end{bmatrix}$ 与向量 $\boldsymbol{y}=\begin{bmatrix}-1\\-5\end{bmatrix}$ 之间的距离.

14. 计算向量 $\boldsymbol{u}=\begin{bmatrix}0\\-5\\2\end{bmatrix}$ 与向量 $\boldsymbol{z}=\begin{bmatrix}-4\\-1\\8\end{bmatrix}$ 之间的距离.

在习题 15~18 中，确定哪一对向量相互正交.

15. $\boldsymbol{a}=\begin{bmatrix}8\\-5\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}-2\\-3\end{bmatrix}$
16. $\boldsymbol{u}=\begin{bmatrix}12\\3\\-5\end{bmatrix}$, $\boldsymbol{v}=\begin{bmatrix}2\\-3\\3\end{bmatrix}$
17. $\boldsymbol{u}=\begin{bmatrix}3\\2\\-5\\0\end{bmatrix}$, $\boldsymbol{v}=\begin{bmatrix}-4\\1\\-2\\6\end{bmatrix}$
18. $\boldsymbol{y}=\begin{bmatrix}-3\\7\\4\\0\end{bmatrix}$, $\boldsymbol{z}=\begin{bmatrix}1\\-8\\15\\-7\end{bmatrix}$

在习题 19~20 中，所有向量在 $\mathbb{R}^n$ 中，说明每个命题的真假，并验证你的答案.

19. a. $\boldsymbol{v}\cdot\boldsymbol{v}=\|\boldsymbol{v}\|^2$.
    b. 对任意数 c， $\boldsymbol{u}\cdot(c\boldsymbol{v})=c(\boldsymbol{u}\cdot\boldsymbol{v})$
    c. 如果向量 $\boldsymbol{u}$ 到向量 $\boldsymbol{v}$ 的距离等于向量 $\boldsymbol{u}$ 到向量 $-\boldsymbol{v}$ 的距离，那么 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是正交的.
    d. 对于一个方阵 A， Col A 中的向量与 Nul A 中的向量正交.

e. 如果向量 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p$ 生成子空间 W，且向量 $\boldsymbol{x}$ 与每一个 $\boldsymbol{v}_i(i=1,\cdots,p)$ 正交，那么向量 $\boldsymbol{x}$ 属于 $W^{\perp}$.

20. a. $\boldsymbol{u}\cdot\boldsymbol{v}-\boldsymbol{v}\cdot\boldsymbol{u}=0$.

b. 对任意数 c，$\|c\boldsymbol{v}\|=c\|\boldsymbol{v}\|$.

c. 如果 $\boldsymbol{x}$ 与子空间 W 中的任一向量正交，那么 $\boldsymbol{x}$ 是 $W^{\perp}$ 中的向量.

d. 如果 $\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2=\|\boldsymbol{u}+\boldsymbol{v}\|^2$，那么 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 相互正交.

e. 对任意 $m\times n$ 矩阵 A，A 的零空间中的向量与 A 的行空间的每个向量正交.

21. 利用内积的转置定义，验证定理 1 中的（b）和（c），注意第 2 章的一些内容.

22. 假若 $\boldsymbol{u}=(u_1,u_2,u_3)$，解释为什么 $\boldsymbol{u}\cdot\boldsymbol{u}\geqslant 0$，什么条件下 $\boldsymbol{u}\cdot\boldsymbol{u}=0$？

23. 假若 $\boldsymbol{u}=\begin{bmatrix}2\\-5\\-1\end{bmatrix}$ 和 $\boldsymbol{v}=\begin{bmatrix}-7\\-4\\6\end{bmatrix}$，计算和比较 $\boldsymbol{u}\cdot\boldsymbol{v}$，$\|\boldsymbol{u}\|^2$，$\|\boldsymbol{v}\|^2$ 和 $\|\boldsymbol{u}+\boldsymbol{v}\|^2$，但不能使用勾股定理.

24. 证明 $\mathbb{R}^n$ 空间中向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的平行四边形法则:

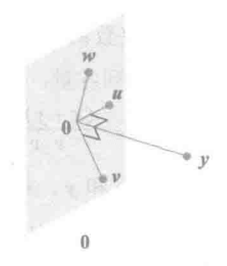
$$\|\boldsymbol{u}+\boldsymbol{v}\|^2+\|\boldsymbol{u}-\boldsymbol{v}\|^2=2\|\boldsymbol{u}\|^2+2\|\boldsymbol{v}\|^2$$

25. 假设 $\boldsymbol{v}=\begin{bmatrix}a\\b\end{bmatrix}$，描述与 $\boldsymbol{v}$ 正交的向量 $\begin{bmatrix}x\\y\end{bmatrix}$ 的集合 H.（提示：考虑 $\boldsymbol{v}=\boldsymbol{0}$, $\boldsymbol{v}\neq\boldsymbol{0}$ 两种情形.）

26. 假设 $\boldsymbol{u}=\begin{bmatrix}5\\-6\\7\end{bmatrix}$，且 W 是 $\mathbb{R}^3$ 中满足 $\boldsymbol{u}\cdot\boldsymbol{x}=0$ 的向量 x 全体，第 4 章中的什么定理可以说明 W 是 $\mathbb{R}^3$ 的子空间？用几何语言描述空间 W.

27. 假若一个向量 $\boldsymbol{y}$ 与向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 都正交，说明 $\boldsymbol{y}$ 与向量 $\boldsymbol{u}+\boldsymbol{v}$ 正交.

28. 假若 $\boldsymbol{y}$ 与向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 都正交，证明 $\boldsymbol{y}$ 与 $\mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$ 中的任一向量正交.（提示：$\mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$ 中的向量 $\boldsymbol{w}$ 具有形式 $\boldsymbol{w}=c_1\boldsymbol{u}+c_2\boldsymbol{v}$. 说明 $\boldsymbol{y}$ 与向量 $\boldsymbol{w}$ 正交.）



0

29. 假若 $W=\mathrm{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$，证明：如果 $\boldsymbol{x}$ 和每个 $\boldsymbol{v}_j$（$1\leqslant j\leqslant p$）正交，那么 $\boldsymbol{x}$ 与 W 中任一向量正交.

30. 假若 W 是 $\mathbb{R}^n$ 的子空间，且 $W^{\perp}$ 是所有与 W 正交的向量集合，利用下列步骤说明 $W^{\perp}$ 是 $\mathbb{R}^n$ 的子空间.

a. 选取 $W^{\perp}$ 中的 $\boldsymbol{z}$，且 $\boldsymbol{u}$ 表示 W 中任意元素，那么 $\boldsymbol{z}\cdot\boldsymbol{u}=0$. 取任意数 c，然后证明 $c\boldsymbol{z}$ 与向量 $\boldsymbol{u}$ 正交（由于 $\boldsymbol{u}$ 是 W 中的任意向量，这说明 $c\boldsymbol{z}$ 在 $W^{\perp}$ 中）.

b. 选取 $\boldsymbol{z}_1$ 和 $\boldsymbol{z}_2$ 在 $W^{\perp}$ 中，且 $\boldsymbol{u}$ 是 W 中任意元素，验证向量 $\boldsymbol{z}_1+\boldsymbol{z}_2$ 与 $\boldsymbol{u}$ 正交，则可从 $\boldsymbol{z}_1+\boldsymbol{z}_2$ 中得出什么结论？为什么？

c. 最后证明 $W^{\perp}$ 是 $\mathbb{R}^n$ 的子空间.

31. 证明：如果 $\boldsymbol{x}$ 是属于空间 W 和 $W^{\perp}$ 的向量，那么 $\boldsymbol{x}=\boldsymbol{0}$.

32. [M] 构造 $\mathbb{R}^4$ 的任意一对向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$，假若：

$$A=\begin{bmatrix}0.5&0.5&0.5&0.5\\0.5&0.5&-0.5&-0.5\\0.5&-0.5&0.5&-0.5\\0.5&-0.5&-0.5&0.5\end{bmatrix}$$

a. 记 A 的列向量为 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_4$，计算每一列的长度和 $\boldsymbol{a}_1\cdot\boldsymbol{a}_2$，$\boldsymbol{a}_1\cdot\boldsymbol{a}_3$，$\boldsymbol{a}_1\cdot\boldsymbol{a}_4$，$\boldsymbol{a}_2\cdot\boldsymbol{a}_3$，$\boldsymbol{a}_2\cdot\boldsymbol{a}_4$，$\boldsymbol{a}_3\cdot\boldsymbol{a}_4$.

b. 计算并比较 $\boldsymbol{u}$ 和 $A\boldsymbol{u}$，$\boldsymbol{v}$ 和 $A\boldsymbol{v}$ 的长度.

c. 利用方程（2）计算向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 夹角的余弦值，并将此值和向量 $A\boldsymbol{u}$ 和 $A\boldsymbol{v}$ 夹角的余弦值相比较.

d. 对任意两个向量，重复步骤（b）和（c），从 A 对任意向量的作用可以得出什么猜想？

33. [M] $\mathbb{R}^4$ 中分量是整数的任意向量 $\boldsymbol{x}$，$\boldsymbol{y}$ 和 $\boldsymbol{v}$（且 $\boldsymbol{v}\neq\boldsymbol{0}$），计算下列各量：
$\left(\frac{\boldsymbol{x}\cdot\boldsymbol{v}}{\boldsymbol{v}\cdot\boldsymbol{v}}\right)\boldsymbol{v}$，$\left(\frac{\boldsymbol{y}\cdot\boldsymbol{v}}{\boldsymbol{v}\cdot\boldsymbol{v}}\right)\boldsymbol{v}$，$\frac{(\boldsymbol{x}+\boldsymbol{y})\cdot\boldsymbol{v}}{\boldsymbol{v}\cdot\boldsymbol{v}}\boldsymbol{v}$，$\frac{(10\boldsymbol{x})\cdot\boldsymbol{v}}{\boldsymbol{v}\cdot\boldsymbol{v}}\boldsymbol{v}$
选取新的任意向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$，重复上面计算，从映射 $\boldsymbol{x}\mapsto T(\boldsymbol{x})=\left(\frac{\boldsymbol{x}\cdot\boldsymbol{v}}{\boldsymbol{v}\cdot\boldsymbol{v}}\right)\boldsymbol{v}$（对 $\boldsymbol{v}\neq\boldsymbol{0}$）可得出什么猜想，用代数证明你的猜想.

34. [M] 设 $A=\begin{bmatrix}-6&3&-27&-33&-13\\6&-5&25&28&14\\8&-6&34&38&18\\12&-10&50&41&23\\14&-21&49&29&33\end{bmatrix}$，构造矩阵 N 使其列是 Nul A 的一组基，构造矩阵 R 使其行是 Row A 的一组基（见 4.6 节）. 用 N 和 R 执行矩阵计算来说明定理 3 的结论.

练习题答案

1. $\boldsymbol{a}\cdot\boldsymbol{b}=7$，$\boldsymbol{a}\cdot\boldsymbol{a}=5$，因此，$\frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\boldsymbol{a}\cdot\boldsymbol{a}}=\frac{7}{5}$，且 $\left(\frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\boldsymbol{a}\cdot\boldsymbol{a}}\right)\boldsymbol{a}=\frac{7}{5}\boldsymbol{a}=\begin{bmatrix}-14/5\\7/5\end{bmatrix}$.

2. 先倍乘 $\boldsymbol{c}$，乘 3 得到 $\boldsymbol{y}=\begin{bmatrix}4\\-3\\2\end{bmatrix}$，计算 $\|\boldsymbol{y}\|^2=29$ 和 $\|\boldsymbol{y}\|=\sqrt{29}$. 与向量 $\boldsymbol{c}$ 和 $\boldsymbol{y}$ 方向一致的单位向量是：

$$\boldsymbol{u}=\frac{1}{\|\boldsymbol{y}\|}\boldsymbol{y}=\begin{bmatrix}4/\sqrt{29}\\-3/\sqrt{29}\\2/\sqrt{29}\end{bmatrix}$$

3. $\boldsymbol{d}$ 与 $\boldsymbol{c}$ 正交，因为：

$$\boldsymbol{d}\cdot\boldsymbol{c}=\begin{bmatrix}5\\6\\-1\end{bmatrix}\cdot\begin{bmatrix}4/3\\-1\\2/3\end{bmatrix}=\frac{20}{3}-6-\frac{2}{3}=0$$

4. $\boldsymbol{d}$ 与 $\boldsymbol{u}$ 正交，因为向量 $\boldsymbol{u}$ 具有形式 $k\boldsymbol{c}$，且对任意 k，

$$\boldsymbol{d}\cdot\boldsymbol{u}=\boldsymbol{d}\cdot(k\boldsymbol{c})=k(\boldsymbol{d}\cdot\boldsymbol{c})=k(0)=0$$

6.2 正交集

$\mathbb{R}^n$ 中的向量集合 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 称为正交向量集，如果集合中的任意两个不同向量都正交，即当 $i\neq j$ 时，$\boldsymbol{u}_i\cdot\boldsymbol{u}_j=0$.

例 1 证明 $\{\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3\}$ 是一个正交集，此处

$$\boldsymbol{u}_1=\begin{bmatrix}3\\1\\1\end{bmatrix},\quad \boldsymbol{u}_2=\begin{bmatrix}-1\\2\\1\end{bmatrix},\quad \boldsymbol{u}_3=\begin{bmatrix}-1/2\\-2\\7/2\end{bmatrix}$$

解 考察三种可能的不同向量对，即 $\{\boldsymbol{u}_1,\boldsymbol{u}_2\},\{\boldsymbol{u}_1,\boldsymbol{u}_3\}$ 和 $\{\boldsymbol{u}_2,\boldsymbol{u}_3\}$.

$$\boldsymbol{u}_1\cdot\boldsymbol{u}_2=3(-1)+1(2)+1(1)=0$$
$$\boldsymbol{u}_1\cdot\boldsymbol{u}_3=3(-1/2)+1(-2)+1(7/2)=0$$
$$\boldsymbol{u}_2\cdot\boldsymbol{u}_3=-1(-1/2)+2(-2)+1(7/2)=0$$

每对不同的向量是垂直的，所以 $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ 是正交集，如图 6-10 中的三条线段，它们之间相互垂直.

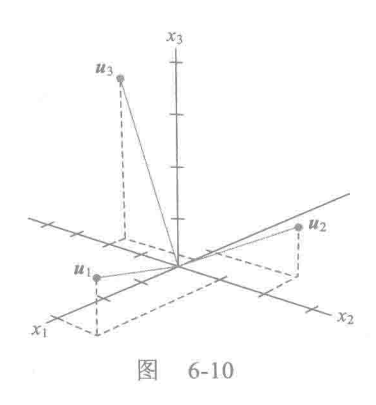


图 6-10 ■

定理 4 如果 $S=\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是由 $\mathbb{R}^n$ 空间中非零向量构成的正交集，那么 S 是线性无关集，因此构成所生成的子空间 S 的一组基.

证 如果 $\boldsymbol{0}=c_1\boldsymbol{u}_1+\cdots+c_p\boldsymbol{u}_p$ 对任意数 $c_1,\cdots,c_p$ 成立，那么

$$\begin{aligned}0=\boldsymbol{0}\cdot\boldsymbol{u}_1&=(c_1\boldsymbol{u}_1+c_2\boldsymbol{u}_2+\cdots+c_p\boldsymbol{u}_p)\cdot\boldsymbol{u}_1\\&=(c_1\boldsymbol{u}_1)\cdot\boldsymbol{u}_1+(c_2\boldsymbol{u}_2)\cdot\boldsymbol{u}_1+\cdots+(c_p\boldsymbol{u}_p)\cdot\boldsymbol{u}_1\\&=c_1(\boldsymbol{u}_1\cdot\boldsymbol{u}_1)+c_2(\boldsymbol{u}_2\cdot\boldsymbol{u}_1)+\cdots+c_p(\boldsymbol{u}_p\cdot\boldsymbol{u}_1)\\&=c_1(\boldsymbol{u}_1\cdot\boldsymbol{u}_1)\end{aligned}$$

因为 $\boldsymbol{u}_1$ 非零，且 $\boldsymbol{u}_1$ 与 $\boldsymbol{u}_2,\cdots,\boldsymbol{u}_p$ 正交，由 $\boldsymbol{u}_1\cdot\boldsymbol{u}_1$ 非零得到 $c_1=0$. 类似可得 $c_2,\cdots,c_p$ 为零，从而得到 S 是线性无关集. ■

定义 $\mathbb{R}^n$ 中子空间 W 的一个正交基是 W 的一个基，且是正交集.

下面的定理表明正交基比其他基优越，线性组合中的权值较易计算.

定理 5 假设 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是 $\mathbb{R}^n$ 中子空间 W 的正交基，对 W 中的每个向量 $\boldsymbol{y}$，线性组合 $\boldsymbol{y}=c_1\boldsymbol{u}_1+\cdots+c_p\boldsymbol{u}_p$ 中的权值可以由 $c_j=\dfrac{\boldsymbol{y}\cdot\boldsymbol{u}_j}{\boldsymbol{u}_j\cdot\boldsymbol{u}_j}\quad(j=1,\cdots,p)$ 计算.

证 像前面证明一样，正交集 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 表明

$$\boldsymbol{y}\cdot\boldsymbol{u}_1=(c_1\boldsymbol{u}_1+c_2\boldsymbol{u}_2+\cdots+c_p\boldsymbol{u}_p)\cdot\boldsymbol{u}_1=c_1(\boldsymbol{u}_1\cdot\boldsymbol{u}_1)$$

由于 $\boldsymbol{u}_1\cdot\boldsymbol{u}_1$ 非零，从上面方程中可以解出系数 c_1，若计算 $\boldsymbol{y}\cdot\boldsymbol{u}_j(j=2,\cdots,p)$，可求出系数 c_j. ■

例 2 例 1 中的集合 $S=\{\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3\}$ 是 $\mathbb{R}^3$ 中的一个正交基，将向量 $\boldsymbol{y}=\begin{bmatrix}6\\1\\-8\end{bmatrix}$ 表示成 S 中向量的线性组合.

解 计算

$$y\cdot u_1=11 \qquad y\cdot u_2=-12 \qquad y\cdot u_3=-33$$
$$u_1\cdot u_1=11 \qquad u_2\cdot u_2=6 \qquad u_3\cdot u_3=33/2$$

由定理 5 得

$$\begin{aligned} y &= \frac{y\cdot u_1}{u_1\cdot u_1}\cdot u_1+\frac{y\cdot u_2}{u_2\cdot u_2}\cdot u_2+\frac{y\cdot u_3}{u_3\cdot u_3}\cdot u_3 \\ &= \frac{11}{11}u_1+\frac{-12}{6}u_2+\frac{-33}{33/2}u_3 \\ &= u_1-2u_2-2u_3 \end{aligned}$$

■

注意到，由正交基构成的线性表示，y 的权值十分容易计算. 如果基不是正交的，则必须类似第 1 章解线性方程组才能得到.

下面我们构造一个非常重要的步骤，涉及许多包含正交计算的问题，而且它会对定理 5 给出一个几何解释.

正交投影

对 $\mathbb{R}^n$ 中给出的非零向量 u，考虑 $\mathbb{R}^n$ 中一个向量 y 分解为两个向量之和的问题，一个向量是向量 u 的数量乘积，另一个向量与 u 垂直. 我们期望写成

$$y=\hat{y}+z \tag{1}$$

其中 $\hat{y}=\alpha u$，α 是一个数，z 是一个垂直于 u 的向量，见图 6-11. 对给定数 α，记 $z=y-\alpha u$，则方程（1）可以满足，那么 $y-\hat{y}$ 和 u 正交的充分必要条件是

$$0=(y-\alpha u)\cdot u=y\cdot u-(\alpha u)\cdot u=y\cdot u-\alpha(u\cdot u)$$

也就是满足方程(1)，且 z 与 u 正交的充分必要条件是 $\alpha=\dfrac{y\cdot u}{u\cdot u}$ 且 $\hat{y}=\dfrac{y\cdot u}{u\cdot u}\cdot u$. 向量 $\hat{y}$ 称为 **y 在 u 上的正交投影**，向量 z 称为 **y 垂直 u 的分量**.

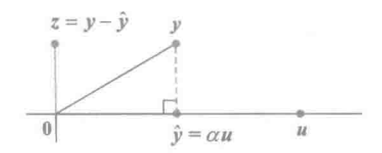


图 6-11 求α使 $y-\hat{y}$ 正交于 u

如果 c 是非零数，且在 $\hat{y}$ 的定义中用 cu 代替 u，那么 y 在 cu 上的正交投影和 y 在 u 上的正交投影完全一致（习题 31），因此这个投影可由 u 向量所生成的子空间 L（经过 u 和原点的直线）所确定. 有时用 $\text{proj}_L\, y$ 来表示 $\hat{y}$，并称之为 **y 在 u 上的正交投影**，即

$$\hat{y}=\text{proj}_L\, y=\frac{y\cdot u}{u\cdot u}\cdot u \tag{2}$$

例 3 假设 $y=\begin{bmatrix}7\\6\end{bmatrix}$ 和 $u=\begin{bmatrix}4\\2\end{bmatrix}$，找出 y 在 u 上的正交投影，然后将 y 写成两个正交向量之和，

一个在 Span$\{\boldsymbol{u}\}$，另一个与 $\boldsymbol{u}$ 正交.

解 计算

$$\boldsymbol{y}\cdot\boldsymbol{u}=\begin{bmatrix}7\\6\end{bmatrix}\cdot\begin{bmatrix}4\\2\end{bmatrix}=40$$

$$\boldsymbol{u}\cdot\boldsymbol{u}=\begin{bmatrix}4\\2\end{bmatrix}\cdot\begin{bmatrix}4\\2\end{bmatrix}=20$$

则 $\boldsymbol{y}$ 在 $\boldsymbol{u}$ 上的正交投影是：

$$\hat{\boldsymbol{y}}=\frac{\boldsymbol{y}\cdot\boldsymbol{u}}{\boldsymbol{u}\cdot\boldsymbol{u}}\cdot\boldsymbol{u}=\frac{40}{20}\boldsymbol{u}=2\begin{bmatrix}4\\2\end{bmatrix}=\begin{bmatrix}8\\4\end{bmatrix}$$

$\boldsymbol{y}$ 垂直于 $\boldsymbol{u}$ 的分量是：

$$\boldsymbol{y}-\hat{\boldsymbol{y}}=\begin{bmatrix}7\\6\end{bmatrix}-\begin{bmatrix}8\\4\end{bmatrix}=\begin{bmatrix}-1\\2\end{bmatrix}$$

两个向量之和为 $\boldsymbol{y}$，即

$$\underset{\boldsymbol{y}}{\begin{bmatrix}7\\6\end{bmatrix}}=\underset{\hat{\boldsymbol{y}}}{\begin{bmatrix}8\\4\end{bmatrix}}+\underset{(\boldsymbol{y}-\hat{\boldsymbol{y}})}{\begin{bmatrix}-1\\2\end{bmatrix}}$$

向量 $\boldsymbol{y}$ 的分解可表示为图 6-12. 注意：如果上面的计算正确，那么 $\{\hat{\boldsymbol{y}},\boldsymbol{y}-\hat{\boldsymbol{y}}\}$ 是正交集. 作为检验，计算

$$\hat{\boldsymbol{y}}\cdot(\boldsymbol{y}-\hat{\boldsymbol{y}})=\begin{bmatrix}8\\4\end{bmatrix}\cdot\begin{bmatrix}-1\\2\end{bmatrix}=-8+8=0$$

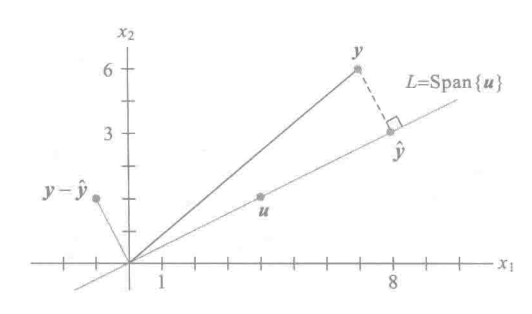


图 6-12 $\hat{\boldsymbol{y}}$ 在通过原点的直线 L 上的正交投影 ■

由于图 6-12 中连接 $\boldsymbol{y}$ 与 $\hat{\boldsymbol{y}}$ 的线段垂直于 L，由 $\hat{\boldsymbol{y}}$ 的构造可知，标记为 $\hat{\boldsymbol{y}}$ 的点是 $\boldsymbol{y}$ 距离 L 的最近点.（这可用几何方法证明，这里我们假设 $\mathbb{R}^2$ 中成立，在 6.3 节给出 $\mathbb{R}^n$ 情形的证明.）

例 4 计算图 6-12 中从 $\boldsymbol{y}$ 到 L 的距离.

解 从 $\boldsymbol{y}$ 到 L 的距离，是从 $\boldsymbol{y}$ 到正交投影垂直线段的长度，这个长度等于 $\boldsymbol{y}-\hat{\boldsymbol{y}}$ 的长度，从而距离为

$$\|\boldsymbol{y}-\hat{\boldsymbol{y}}\|=\sqrt{(-1)^2+2^2}=\sqrt{5}$$ ■

定理 5 的几何解释

(2) 式中正交投影 $\hat{\boldsymbol{y}}$ 的公式和定理 5 中每一项形式一致，这样，定理 5 将向量 $\boldsymbol{y}$ 分解为一维子空间上正交投影之和.

对于 $W=\mathbb{R}^2=\text{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$， $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 相互正交的情形，很容易看到分解式，对任意 $\mathbb{R}^2$ 中的向量 $\boldsymbol{y}$ 可以写成：

$$\boldsymbol{y}=\frac{\boldsymbol{y}\cdot\boldsymbol{u}_1}{\boldsymbol{u}_1\cdot\boldsymbol{u}_1}\boldsymbol{u}_1+\frac{\boldsymbol{y}\cdot\boldsymbol{u}_2}{\boldsymbol{u}_2\cdot\boldsymbol{u}_2}\boldsymbol{u}_2 \tag{3}$$

(3) 中的第一项是 $\boldsymbol{y}$ 在子空间 $\text{Span}\{\boldsymbol{u}_1\}$ 上的投影（通过原点和 $\boldsymbol{u}_1$ 的直线），第二项 $\boldsymbol{y}$ 在子空间 $\text{Span}\{\boldsymbol{u}_2\}$ 上的投影，(3) 式将 $\boldsymbol{y}$ 表示为由 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 确定的（正交）轴上的投影之和，见图 6-13 .

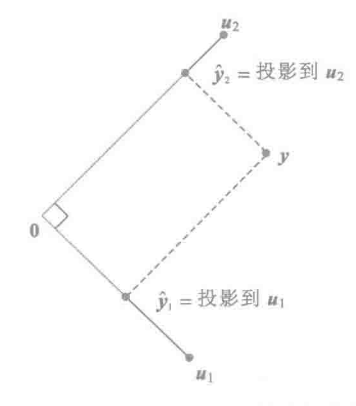


图 6-13 一个向量分解为两个投影之和

定理 5 将 $\text{Span}\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 中的每一个 $\boldsymbol{y}$ 分解成 p 个相互正交的一维子空间上的投影之和.

一个力分解为力的分量

如果某一力施加到一个物体上，在物理上可能出现图 6-13 的分解，通过选取合适的坐标系统，一个力可以表示为 $\mathbb{R}^2$ 或 $\mathbb{R}^3$ 中向量 $\boldsymbol{y}$，常常这类问题包含某些特别感兴趣的方向，可表示为另一个向量 $\boldsymbol{u}$. 例如，一个在水平方向移动的物体被施加外力后，用向量 $\boldsymbol{u}$ 表示移动的方向，见图 6-14. 一个关键问题是将力分解为 $\boldsymbol{u}$ 方向的分量和与 $\boldsymbol{u}$ 正交方向的分量，具体计算类似于上面已完成的例 3.

图 6-14

单位正交集

集合 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是一个**单位正交集**，如果它是由单位向量构成的正交集. 如果 W 是一个由单位正交集合组成的子空间，那么 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是 W 的**单位正交基**，原因是这类集合自然线性无关，见定理 4.

最简单的单位正交集合是 $\mathbb{R}^n$ 中的标准基 $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$，任何集合 $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$ 的非空子集也是单位正交的，下面是一个更复杂的例子.

例 5 证明 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 是 $\mathbb{R}^3$ 的一个单位正交基，其中：

$$\boldsymbol{v}_1=\begin{bmatrix}3/\sqrt{11}\\1/\sqrt{11}\\1/\sqrt{11}\end{bmatrix}\qquad \boldsymbol{v}_2=\begin{bmatrix}-1/\sqrt{6}\\2/\sqrt{6}\\1/\sqrt{6}\end{bmatrix}\qquad \boldsymbol{v}_3=\begin{bmatrix}-1/\sqrt{66}\\-4/\sqrt{66}\\7/\sqrt{66}\end{bmatrix}$$

解 计算

$$\boldsymbol{v}_1\cdot\boldsymbol{v}_2=-3/\sqrt{66}+2/\sqrt{66}+1/\sqrt{66}=0$$
$$\boldsymbol{v}_1\cdot\boldsymbol{v}_3=-3/\sqrt{726}-4/\sqrt{726}+7/\sqrt{726}=0$$
$$\boldsymbol{v}_2\cdot\boldsymbol{v}_3=1/\sqrt{396}-8/\sqrt{396}+7/\sqrt{396}=0$$

从而 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 是一个正交基，另外

$$\boldsymbol{v}_1\cdot\boldsymbol{v}_1=9/11+1/11+1/11=1$$
$$\boldsymbol{v}_2\cdot\boldsymbol{v}_2=1/6+4/6+1/6=1$$
$$\boldsymbol{v}_3\cdot\boldsymbol{v}_3=1/66+16/66+49/66=1$$

从而证明 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 和 $\boldsymbol{v}_3$ 是单位向量，即 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 是一个单位正交集. 由于集合线性无关，它的三个向量构成 $\mathbb{R}^3$ 的一个基，见图 6-15.

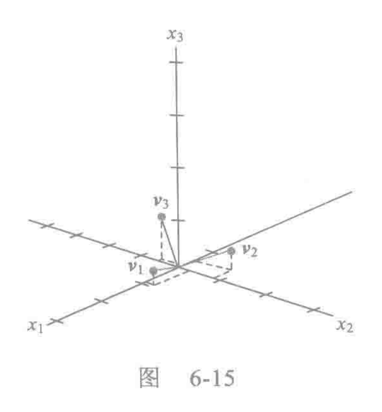


图 6-15 ■

当一个正交集中的向量被“单位化”具有单位长度后，这些新向量仍然保持正交性，因此新的集合成为单位正交基，见习题 32. 非常容易检查图 6-15 中的向量（例 5）是图 6-10（例 1）中向量各个方向的单位化向量.

各列形成单位正交基的矩阵，在应用和矩阵算法的计算中都非常重要，它们的主要性质由下面定理 6 和定理 7 给出.

定理 6 一个 $m\times n$ 矩阵 U 具有单位正交列向量的充分必要条件是 $U^{\mathrm{T}}U=I$.

证 为简化运算，我们假设 U 仅有三列，每列都是 $\mathbb{R}^m$ 中的向量，一般情形的证明本质上完全一致. 设 $U=[\boldsymbol{u}_1\ \ \boldsymbol{u}_2\ \ \boldsymbol{u}_3]$，并且计算：

$$U^{\mathrm{T}}U=\begin{bmatrix}\boldsymbol{u}_1^{\mathrm{T}}\\ \boldsymbol{u}_2^{\mathrm{T}}\\ \boldsymbol{u}_3^{\mathrm{T}}\end{bmatrix}[\boldsymbol{u}_1\ \ \boldsymbol{u}_2\ \ \boldsymbol{u}_3]=\begin{bmatrix}\boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{u}_1 & \boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{u}_2 & \boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{u}_3\\ \boldsymbol{u}_2^{\mathrm{T}}\boldsymbol{u}_1 & \boldsymbol{u}_2^{\mathrm{T}}\boldsymbol{u}_2 & \boldsymbol{u}_2^{\mathrm{T}}\boldsymbol{u}_3\\ \boldsymbol{u}_3^{\mathrm{T}}\boldsymbol{u}_1 & \boldsymbol{u}_3^{\mathrm{T}}\boldsymbol{u}_2 & \boldsymbol{u}_3^{\mathrm{T}}\boldsymbol{u}_3\end{bmatrix} \tag{4}$$

右边矩阵中的元素是利用转置 T 表示的内积，U 的列向量是正交的充分必要条件是

$$\boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{u}_2=\boldsymbol{u}_2^{\mathrm{T}}\boldsymbol{u}_1=0,\quad \boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{u}_3=\boldsymbol{u}_3^{\mathrm{T}}\boldsymbol{u}_1=0,\quad \boldsymbol{u}_2^{\mathrm{T}}\boldsymbol{u}_3=\boldsymbol{u}_3^{\mathrm{T}}\boldsymbol{u}_2=0 \tag{5}$$

U 的列向量是单位长度的充分必要条件是：

$$\boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{u}_1=1,\quad \boldsymbol{u}_2^{\mathrm{T}}\boldsymbol{u}_2=1,\quad \boldsymbol{u}_3^{\mathrm{T}}\boldsymbol{u}_3=1 \tag{6}$$

定理可以从（4）~（6）立刻得出. ■

定理 7 假设 U 是一个具有单位正交列的 $m\times n$ 矩阵，且 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 是 $\mathbb{R}^n$ 的向量，那么

a. $\|U\boldsymbol{x}\|=\|\boldsymbol{x}\|$.

b. $(U\boldsymbol{x})\cdot(U\boldsymbol{y})=\boldsymbol{x}\cdot\boldsymbol{y}$.

c. $(U\boldsymbol{x})\cdot(U\boldsymbol{y})=0$ 的充分必要条件是 $\boldsymbol{x}\cdot\boldsymbol{y}=0$.

性质（a）和（c）表明，线性映射 $\boldsymbol{x}\mapsto U\boldsymbol{x}$ 保持长度和正交性，这个性质对很多计算机算法非常重要，定理 7 的具体证明见习题 25.

例 6 若 $U=\begin{bmatrix}1/\sqrt{2} & 2/3\\ 1/\sqrt{2} & -2/3\\ 0 & 1/3\end{bmatrix}$ 和 $\boldsymbol{x}=\begin{bmatrix}\sqrt{2}\\ 3\end{bmatrix}$，注意 U 具有单位正交列，并且

$$U^{\mathrm{T}}U=\begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2} & 0\\ 2/3 & -2/3 & 1/3\end{bmatrix}\begin{bmatrix}1/\sqrt{2} & 2/3\\ 1/\sqrt{2} & -2/3\\ 0 & 1/3\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

验证 $\|U\boldsymbol{x}\|=\|\boldsymbol{x}\|$.

解

$$U\boldsymbol{x}=\begin{bmatrix}1/\sqrt{2} & 2/3\\ 1/\sqrt{2} & -2/3\\ 0 & 1/3\end{bmatrix}\begin{bmatrix}\sqrt{2}\\ 3\end{bmatrix}=\begin{bmatrix}3\\ -1\\ 1\end{bmatrix}$$

$$\|U\boldsymbol{x}\|=\sqrt{9+1+1}=\sqrt{11},\quad \|\boldsymbol{x}\|=\sqrt{2+9}=\sqrt{11}$$ ■

当矩阵是方阵时，定理 6 和定理 7 非常有用. 一个**正交矩阵**就是一个可逆的方阵 U，且满足 $U^{-1}=U^{\mathrm{T}}$，由定理 6，这样的矩阵具有单位正交列.[㊀] 很容易验证，任何列单位正交的方阵是正交矩阵，恰巧，这类矩阵同样具有单位正交的行，见习题 27 和习题 28. 正交矩阵在第 7 章有广泛的应用.

㊀ 一个更好的术语应该是标准正交矩阵，这个术语可以在一些统计教材中找到. 但是，正交矩阵是线性代数中的标准术语.

例 7 矩阵

$$U=\begin{bmatrix}3/\sqrt{11} & -1/\sqrt{6} & -1/\sqrt{66}\\ 1/\sqrt{11} & 2/\sqrt{6} & -4/\sqrt{66}\\ 1/\sqrt{11} & 1/\sqrt{6} & 7/\sqrt{66}\end{bmatrix}$$

是单位正交矩阵，因为它是方阵且它的列是单位正交的，见例 5，事实上，它的行也是单位正交的. ■

练习题

1. 设 $u_1=\begin{bmatrix}-1/\sqrt{5}\\ 2/\sqrt{5}\end{bmatrix}$， $u_2=\begin{bmatrix}2/\sqrt{5}\\ 1/\sqrt{5}\end{bmatrix}$，说明 $\{u_1,u_2\}$ 是 $\mathbb{R}^2$ 的单位正交基.

2. 设 y 和 L 是例 3 和图 6-12 中的记号，计算 y 在 L 上的正交投影 $\hat{y}$，用 $u=\begin{bmatrix}2\\1\end{bmatrix}$ 替换例 3 中的 u.

3. 设 U 和 x 是例 6 中的记号，且 $y=\begin{bmatrix}-3\sqrt{2}\\ 6\end{bmatrix}$. 验证：$Ux\cdot Uy=x\cdot y$.

习题 6.2

在习题 1~6 中，判断哪一个向量的集合是正交的.

1. $\begin{bmatrix}-1\\4\\-3\end{bmatrix}$, $\begin{bmatrix}5\\2\\1\end{bmatrix}$, $\begin{bmatrix}3\\-4\\-7\end{bmatrix}$

2. $\begin{bmatrix}1\\-2\\1\end{bmatrix}$, $\begin{bmatrix}0\\1\\2\end{bmatrix}$, $\begin{bmatrix}-5\\-2\\1\end{bmatrix}$

3. $\begin{bmatrix}2\\-7\\-1\end{bmatrix}$, $\begin{bmatrix}-6\\-3\\9\end{bmatrix}$, $\begin{bmatrix}3\\1\\-1\end{bmatrix}$

4. $\begin{bmatrix}2\\-5\\-3\end{bmatrix}$, $\begin{bmatrix}0\\0\\0\end{bmatrix}$, $\begin{bmatrix}4\\-2\\6\end{bmatrix}$

5. $\begin{bmatrix}3\\-2\\1\\3\end{bmatrix}$, $\begin{bmatrix}-1\\3\\-3\\4\end{bmatrix}$, $\begin{bmatrix}3\\8\\7\\0\end{bmatrix}$

6. $\begin{bmatrix}5\\-4\\0\\3\end{bmatrix}$, $\begin{bmatrix}-4\\1\\-3\\8\end{bmatrix}$, $\begin{bmatrix}3\\3\\5\\-1\end{bmatrix}$

在习题 7~10 中，证明 $\{u_1,u_2\}$ 或者 $\{u_1,u_2,u_3\}$ 分别是 $\mathbb{R}^2$ 和 $\mathbb{R}^3$ 的正交基，且将 x 表示为这些 u 的线性组合.

7. $u_1=\begin{bmatrix}2\\-3\end{bmatrix}$， $u_2=\begin{bmatrix}6\\4\end{bmatrix}$， $x=\begin{bmatrix}9\\-7\end{bmatrix}$

8. $u_1=\begin{bmatrix}3\\1\end{bmatrix}$， $u_2=\begin{bmatrix}-2\\6\end{bmatrix}$， $x=\begin{bmatrix}-6\\3\end{bmatrix}$

9. $u_1=\begin{bmatrix}1\\0\\1\end{bmatrix}$， $u_2=\begin{bmatrix}-1\\4\\1\end{bmatrix}$， $u_3=\begin{bmatrix}2\\1\\-2\end{bmatrix}$， $x=\begin{bmatrix}8\\-4\\-3\end{bmatrix}$

10. $u_1=\begin{bmatrix}3\\-3\\0\end{bmatrix}$， $u_2=\begin{bmatrix}2\\2\\-1\end{bmatrix}$， $u_3=\begin{bmatrix}1\\1\\4\end{bmatrix}$， $x=\begin{bmatrix}5\\-3\\1\end{bmatrix}$

11. 计算向量 $\begin{bmatrix}1\\7\end{bmatrix}$ 在通过 $\begin{bmatrix}-4\\2\end{bmatrix}$ 和原点的直线上的正交投影.

12. 计算向量 $\begin{bmatrix}1\\-1\end{bmatrix}$ 在通过 $\begin{bmatrix}-1\\3\end{bmatrix}$ 和原点的直线上的正交投影.

13. 若 $y=\begin{bmatrix}2\\3\end{bmatrix}$ 和 $u=\begin{bmatrix}4\\-7\end{bmatrix}$，将 y 写成两个正交向量之和，一个属于 $\mathrm{Span}\{u\}$，另一个与 u 正交.

14. 若 $y=\begin{bmatrix}2\\6\end{bmatrix}$ 和 $u=\begin{bmatrix}7\\1\end{bmatrix}$，将 y 写成两个正交向量之和，一个属于 $\mathrm{Span}\{u\}$，另一个与 u 正交.

15. 若 $y=\begin{bmatrix}3\\1\end{bmatrix}$ 和 $u=\begin{bmatrix}8\\6\end{bmatrix}$，计算向量 y 与通过 u 和原点的直线之间的距离.

16. 若 $y=\begin{bmatrix}-3\\9\end{bmatrix}$ 和 $u=\begin{bmatrix}1\\2\end{bmatrix}$，计算向量 y 与通过 u 和原点的直线之间的距离.

在习题 17~22 中，确定哪一个向量集合是正交的，如果集合只是正交的，将向量单位化产生一个单位正交集.

17. $\begin{bmatrix}1/3\\1/3\\1/3\end{bmatrix}$，$\begin{bmatrix}-1/2\\0\\1/2\end{bmatrix}$　　18. $\begin{bmatrix}0\\1\\0\end{bmatrix}$，$\begin{bmatrix}0\\-1\\0\end{bmatrix}$

19. $\begin{bmatrix}-0.6\\0.8\end{bmatrix}$，$\begin{bmatrix}0.8\\0.6\end{bmatrix}$　　20. $\begin{bmatrix}-2/3\\1/3\\2/3\end{bmatrix}$，$\begin{bmatrix}1/3\\2/3\\0\end{bmatrix}$

21. $\begin{bmatrix}1/\sqrt{10}\\3/\sqrt{20}\\3/\sqrt{20}\end{bmatrix}$，$\begin{bmatrix}3/\sqrt{10}\\-1/\sqrt{20}\\-1/\sqrt{20}\end{bmatrix}$，$\begin{bmatrix}0\\-1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}$

22. $\begin{bmatrix}1/\sqrt{18}\\4/\sqrt{18}\\1/\sqrt{18}\end{bmatrix}$，$\begin{bmatrix}1/\sqrt{2}\\0\\-1/\sqrt{2}\end{bmatrix}$，$\begin{bmatrix}-2/3\\1/3\\-2/3\end{bmatrix}$

在习题 23~24 中，所有向量属于 $\mathbb{R}^n$，判断下述论断的正误，验证每一个结论.

23. a. $\mathbb{R}^n$ 上的每一个线性无关集并非都是正交集.
  b. 如果 $\boldsymbol{y}$ 是正交集中非零向量的线性组合，那么线性组合的权数可不用矩阵的行变换求得.
  c. 如果非零向量构成的正交集中的向量被单位化，那么，其中一些新向量可能不正交.
  d. 一个具有单位正交列的矩阵是正交矩阵.
  e. 如果 L 是通过原点的直线，并且 $\hat{\boldsymbol{y}}$ 是 $\boldsymbol{y}$ 在 L 上的正交投影，那么 $\|\hat{\boldsymbol{y}}\|$ 表示 $\boldsymbol{y}$ 到 L 的距离.

24. a. $\mathbb{R}^n$ 中的每一个正交集并非都是线性无关的.
  b. 如果集合 $S=\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 具有性质，当 $i\neq j$ 时，$\boldsymbol{u}_i\cdot\boldsymbol{u}_j=0$，那么 S 是单位正交集.
  c. 如果 $m\times n$ 矩阵 A 的列是单位正交，那么线性映射 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 保持长度.
  d. 向量 $\boldsymbol{y}$ 在 $\boldsymbol{v}$ 上的正交投影和 $\boldsymbol{y}$ 在 $c\boldsymbol{v}$（$c\neq 0$）上的正交投影一致.
  e. 正交矩阵是可逆的.

25. 证明定理 7（提示：对（a），计算 $\|U\boldsymbol{x}\|^2$ 或首先证明（b））.

26. 若 W 是 $\mathbb{R}^n$ 中 n 个非零正交向量张成的子空间，试说明 $W=\mathbb{R}^n$.

27. 若 U 是具有单位正交列的方阵，说明 U 为什么可逆.（注意证明中的定理.）

28. 若 U 是 $n\times n$ 正交矩阵，证明 U 的行向量构成 $\mathbb{R}^n$ 的单位正交基.

29. 若 U 和 V 是正交矩阵，说明为什么 UV 也是正交矩阵（即说明为什么 UV 可逆且它的逆为 $(UV)^{\mathrm{T}}$）.

30. 若 U 是正交矩阵且矩阵 V 是交换 U 中某些列得到的矩阵，说明为什么 V 是正交矩阵.

31. 证明：向量 $\boldsymbol{y}$ 在 $\mathbb{R}^2$ 中通过原点的直线 L 上的正交投影 $\hat{\boldsymbol{y}}$ 的公式，不依赖 L 中非零向量 $\boldsymbol{u}$ 的选择，为验证结论，可假设 $\boldsymbol{y}$ 和 $\boldsymbol{u}$ 是给定的，且 $\hat{\boldsymbol{y}}$ 用公式（2）计算，将公式中的 $\boldsymbol{u}$ 用 $c\boldsymbol{u}$ 代替，其中 c 是任意非零值，证明新公式给出同样的 $\hat{\boldsymbol{y}}$.

32. 设 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是非零向量的正交集，c_1 和 c_2 是任意非零数，证明 $\{c_1\boldsymbol{v}_1,c_2\boldsymbol{v}_2\}$ 也是正交基，由于集合的正交性可由成对向量确定，这说明如果正交集中的向量被单位化，新的向量集合仍然是正交的.

33. 设 $\mathbb{R}^n$ 中 $\boldsymbol{u}\neq\boldsymbol{0}$，设 $L=\mathrm{Span}\{\boldsymbol{u}\}$. 证明映射 $\boldsymbol{x}\to\mathrm{proj}_L\boldsymbol{x}$ 是一个线性变换.

34. 设 $\mathbb{R}^n$ 中 $\boldsymbol{u}\neq\boldsymbol{0}$，$L=\mathrm{Span}\{\boldsymbol{u}\}$. 对 $\mathbb{R}^n$ 上的 $\boldsymbol{y}$，**$\boldsymbol{y}$ 在 L 上的反射**是点 $\mathrm{refl}_L\boldsymbol{y}$，定义为 $\mathrm{refl}_L\boldsymbol{y}=2\cdot\mathrm{proj}_L\boldsymbol{y}-\boldsymbol{y}$. 观察图像，其表明 $\mathrm{refl}_L\boldsymbol{y}$ 是 $\hat{\boldsymbol{y}}=\mathrm{proj}_L\boldsymbol{y}$ 与 $\hat{\boldsymbol{y}}-\boldsymbol{y}$ 的和. 见图 6-16. 证明映射 $\boldsymbol{y}\mapsto\mathrm{refl}_L\boldsymbol{y}$ 是一个线性变换.

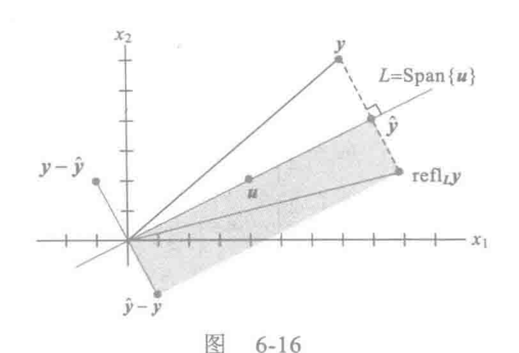


图 6-16

35. [M]证明矩阵 A 的列正交性可由适当的矩阵运算来完成，说明你的计算过程，其中矩阵 A 是：

$$A=\begin{bmatrix}-6&-3&6&1\\-1&2&1&-6\\3&6&3&-2\\6&-3&6&-1\\2&-1&2&3\\-3&6&3&2\\-2&-1&2&-3\\1&2&1&6\end{bmatrix}$$

36. [M] 假若 U 是由习题 35 中矩阵 A 的每一列单位化得到：

a. 计算 U^TU 和 UU^T，它们的不同在哪里？

b. 任意产生一个 $\mathbb{R}^8$ 中的向量 $\boldsymbol{y}$，并且计算 $\boldsymbol{p}=UU^T\boldsymbol{y}$ 和 $\boldsymbol{z}=\boldsymbol{y}-\boldsymbol{p}$，解释为什么 $\boldsymbol{p}$ 属于 Col A，验证 $\boldsymbol{z}$ 和 $\boldsymbol{p}$ 正交.

c. 验证 $\boldsymbol{z}$ 与 U 中每一列正交.

d. 注意 $\boldsymbol{y}=\boldsymbol{p}+\boldsymbol{z}$，$\boldsymbol{p}$ 属于 Col A，解释为什么 $\boldsymbol{z}$ 属于 $(\text{Col } A)^{\perp}$（$\boldsymbol{y}$ 的这个分解的特点将在下一节解释）.

练习题答案

1. 向量相互垂直，因为 $\boldsymbol{u}_1\cdot\boldsymbol{u}_2=-2/5+2/5=0$.

   它们是单位向量，因为

$$\|\boldsymbol{u}_1\|^2=(-1/\sqrt{5})^2+(2/\sqrt{5})^2=1/5+4/5=1$$
$$\|\boldsymbol{u}_2\|^2=(2/\sqrt{5})^2+(1/\sqrt{5})^2=4/5+1/5=1$$

   特别地，集合 $\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$ 线性无关，由于集合包含两个向量，因此构成 $\mathbb{R}^2$ 的一个基.

2. 当 $\boldsymbol{y}=\begin{bmatrix}7\\6\end{bmatrix}$ 和 $\boldsymbol{u}=\begin{bmatrix}2\\1\end{bmatrix}$ 时，有 $\hat{\boldsymbol{y}}=\dfrac{\boldsymbol{y}\cdot\boldsymbol{u}}{\boldsymbol{u}\cdot\boldsymbol{u}}\boldsymbol{u}=\dfrac{20}{5}\begin{bmatrix}2\\1\end{bmatrix}=4\begin{bmatrix}2\\1\end{bmatrix}=\begin{bmatrix}8\\4\end{bmatrix}$.

   这和例 3 中的 $\hat{\boldsymbol{y}}$ 一致，正交投影似乎不依赖于直线上向量 $\boldsymbol{u}$ 的选取，见习题 31.

3. $U\boldsymbol{y}=\begin{bmatrix}1/\sqrt{2}&2/3\\1/\sqrt{2}&-2/3\\0&1/3\end{bmatrix}\begin{bmatrix}-3\sqrt{2}\\6\end{bmatrix}=\begin{bmatrix}1\\-7\\2\end{bmatrix}$.

   同样，由例 6，$\boldsymbol{x}=\begin{bmatrix}\sqrt{2}\\3\end{bmatrix}$ 和 $U\boldsymbol{x}=\begin{bmatrix}3\\-1\\1\end{bmatrix}$. 因此，$U\boldsymbol{x}\cdot U\boldsymbol{y}=3+7+2=12$ 且 $\boldsymbol{x}\cdot\boldsymbol{y}=-6+18=12$.

6.3 正交投影

$\mathbb{R}^2$ 中点在通过原点的直线上的正交投影和空间 $\mathbb{R}^n$ 的情形非常类似，对给定向量 $\boldsymbol{y}$ 和 $\mathbb{R}^n$ 中子空间 W，存在属于 W 的向量 $\hat{\boldsymbol{y}}$ 满足：（1）W 中有唯一向量 $\hat{\boldsymbol{y}}$，使得 $\boldsymbol{y}-\hat{\boldsymbol{y}}$ 与 W 正交，（2）$\hat{\boldsymbol{y}}$ 是 W 中唯一最接近 $\boldsymbol{y}$ 的向量，见图 6-17. $\hat{\boldsymbol{y}}$ 的这两个性质提供了本章介绍实例中提到的求线性方程组的最小二乘解的方法，完整的内容在 6.5 节中学习.

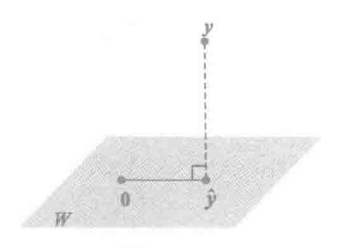


图 6-17

为准备第一个定理，我们注意到，当 $\boldsymbol{y}$ 表示成 $\mathbb{R}^n$ 空间中基 $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n$ 的线性组合时，$\boldsymbol{y}$ 中和的各项可分为两部分，使得 $\boldsymbol{y}$ 写成

$$\boldsymbol{y}=\boldsymbol{z}_1+\boldsymbol{z}_2$$

此处，$\boldsymbol{z}_1$ 是其中一些 $\boldsymbol{u}_i$ 的线性组合，$\boldsymbol{z}_2$ 是其余 $\boldsymbol{u}_i$ 的线性组合，当 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n\}$ 是正交基时，这个思路特别有用. 回忆在 6.1 节，$W^{\perp}$ 表示所有与 W 正交的向量集合.

例 1 假若 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_5\}$ 是 $\mathbb{R}^5$ 中的正交基，且

$$\boldsymbol{y}=c_1\boldsymbol{u}_1+\cdots+c_5\boldsymbol{u}_5$$

考虑子空间 $W=\mathrm{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$，并将 $\boldsymbol{y}$ 写成 W 中向量 $\boldsymbol{z}_1$ 与 $W^{\perp}$ 中向量 $\boldsymbol{z}_2$ 的和.

解 写出分解式

$$\boldsymbol{y}=\underbrace{c_1\boldsymbol{u}_1+c_2\boldsymbol{u}_2}_{\boldsymbol{z}_1}+\underbrace{c_3\boldsymbol{u}_3+c_4\boldsymbol{u}_4+c_5\boldsymbol{u}_5}_{\boldsymbol{z}_2}$$

其中

$$\boldsymbol{z}_1=c_1\boldsymbol{u}_1+c_2\boldsymbol{u}_2\ \text{属于}\ \mathrm{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$$

$$\boldsymbol{z}_2=c_3\boldsymbol{u}_3+c_4\boldsymbol{u}_4+c_5\boldsymbol{u}_5\ \text{属于}\ \mathrm{Span}\{\boldsymbol{u}_3,\boldsymbol{u}_4,\boldsymbol{u}_5\}$$

为证明 $\boldsymbol{z}_2$ 属于 $W^{\perp}$，只需证明 $\boldsymbol{z}_2$ 与以 $\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$ 为基的空间 W 中的向量正交（见 6.1 节），利用内积的性质计算

$$\begin{aligned}\boldsymbol{z}_2\cdot\boldsymbol{u}_1&=(c_3\boldsymbol{u}_3+c_4\boldsymbol{u}_4+c_5\boldsymbol{u}_5)\cdot\boldsymbol{u}_1\\&=c_3\boldsymbol{u}_3\cdot\boldsymbol{u}_1+c_4\boldsymbol{u}_4\cdot\boldsymbol{u}_1+c_5\boldsymbol{u}_5\cdot\boldsymbol{u}_1\\&=0\end{aligned}$$

上式中利用了 $\boldsymbol{u}_1$ 与 $\boldsymbol{u}_3$，$\boldsymbol{u}_4$ 和 $\boldsymbol{u}_5$ 的正交性. 同样的计算可以证明 $\boldsymbol{z}_2\cdot\boldsymbol{u}_2=0$，从而说明 $\boldsymbol{z}_2$ 属于 $W^{\perp}$. ■

下面的定理表明，在没有 $\mathbb{R}^n$ 中的正交基时，对例 1 中的分解 $\boldsymbol{y}=\boldsymbol{z}_1+\boldsymbol{z}_2$ 同样可以计算，只需 W 中有一个正交基即可.

定理 8 （正交分解定理）

若 W 是 $\mathbb{R}^n$ 的一个子空间，那么 $\mathbb{R}^n$ 中每一个向量 $\boldsymbol{y}$ 可以唯一表示

$$\boldsymbol{y}=\hat{\boldsymbol{y}}+\boldsymbol{z}\tag{1}$$

此处 $\hat{\boldsymbol{y}}$ 属于 W 且 $\boldsymbol{z}$ 属于 $W^{\perp}$，实际上，如果 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是 W 的任意正交基，那么

$$\hat{\boldsymbol{y}}=\frac{\boldsymbol{y}\cdot\boldsymbol{u}_1}{\boldsymbol{u}_1\cdot\boldsymbol{u}_1}\boldsymbol{u}_1+\cdots+\frac{\boldsymbol{y}\cdot\boldsymbol{u}_p}{\boldsymbol{u}_p\cdot\boldsymbol{u}_p}\boldsymbol{u}_p\tag{2}$$

且 $\boldsymbol{z}=\boldsymbol{y}-\hat{\boldsymbol{y}}$.

(1) 式中的 $\hat{\boldsymbol{y}}$ 称为 **$\boldsymbol{y}$ 在 W 上的正交投影**，常记作 $\mathrm{proj}_W\,\boldsymbol{y}$. 见图 6-18，当 W 是一维子空间时，$\hat{\boldsymbol{y}}$ 的公式和 6.2 节中的公式一致.

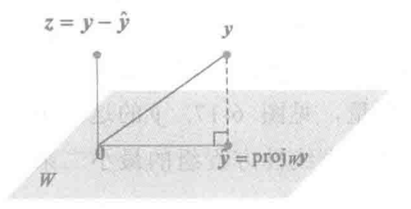


图 6-18 $\boldsymbol{y}$ 在 W 上的正交投影

证 若$\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$是$W$的正交基，且$\hat{\boldsymbol{y}}$用（2）式定义.[㊀] 由于$\hat{\boldsymbol{y}}$是基$\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p$的线性组合，所以$\hat{\boldsymbol{y}}$属于$W$. 假若$\boldsymbol{z}=\boldsymbol{y}-\hat{\boldsymbol{y}}$，由于$\boldsymbol{u}_1$正交于$\boldsymbol{u}_2,\cdots,\boldsymbol{u}_p$，从（2）式可以得出

$$\boldsymbol{z}\cdot\boldsymbol{u}_1=(\boldsymbol{y}-\hat{\boldsymbol{y}})\cdot\boldsymbol{u}_1=\boldsymbol{y}\cdot\boldsymbol{u}_1-\left(\frac{\boldsymbol{y}\cdot\boldsymbol{u}_1}{\boldsymbol{u}_1\cdot\boldsymbol{u}_1}\right)\boldsymbol{u}_1\cdot\boldsymbol{u}_1-0-\cdots-0=\boldsymbol{y}\cdot\boldsymbol{u}_1-\boldsymbol{y}\cdot\boldsymbol{u}_1=0$$

从而$\boldsymbol{z}$与$\boldsymbol{u}_1$正交. 类似地，$\boldsymbol{z}$与空间W的每个基向量$\boldsymbol{u}_j$正交，因此$\boldsymbol{z}$与W中任何向量正交，即$\boldsymbol{z}$属于$W^{\perp}$.

为证明分解（1）是唯一的，若$\boldsymbol{y}$也可以写成$\boldsymbol{y}=\hat{\boldsymbol{y}}_1+\boldsymbol{z}_1$，其中$\hat{\boldsymbol{y}}_1$属于$W$且$\boldsymbol{z}_1$属于$W^{\perp}$，那么$\hat{\boldsymbol{y}}+\boldsymbol{z}=\hat{\boldsymbol{y}}_1+\boldsymbol{z}_1$（由于两边的$\boldsymbol{y}$相同），从而得出

$$\hat{\boldsymbol{y}}-\hat{\boldsymbol{y}}_1=\boldsymbol{z}_1-\boldsymbol{z}$$

这个等式表明向量$\boldsymbol{v}=\hat{\boldsymbol{y}}-\hat{\boldsymbol{y}}_1$属于$W$，又属于$W^{\perp}$（由于$\boldsymbol{z}_1$和$\boldsymbol{z}$属于$W^{\perp}$且$W^{\perp}$是子空间）. 因此$\boldsymbol{v}\cdot\boldsymbol{v}=0$，从而说明$\boldsymbol{v}=\boldsymbol{0}$，即证明$\hat{\boldsymbol{y}}=\hat{\boldsymbol{y}}_1$且$\boldsymbol{z}_1=\boldsymbol{z}$. ■

分解式（1）的唯一性表明，正交投影$\hat{\boldsymbol{y}}$仅依赖于W，并不依赖于（2）中使用的特殊基.

例 2 假设$\boldsymbol{u}_1=\begin{bmatrix}2\\5\\-1\end{bmatrix},\boldsymbol{u}_2=\begin{bmatrix}-2\\1\\1\end{bmatrix}$和$\boldsymbol{y}=\begin{bmatrix}1\\2\\3\end{bmatrix}$. 注意到$\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$是$W=\mathrm{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$的正交基，将$\boldsymbol{y}$写成属于$W$的向量与正交于$W$的向量之和.

解 $\boldsymbol{y}$在W上的正交投影是

$$\begin{aligned}\hat{\boldsymbol{y}}&=\frac{\boldsymbol{y}\cdot\boldsymbol{u}_1}{\boldsymbol{u}_1\cdot\boldsymbol{u}_1}\boldsymbol{u}_1+\frac{\boldsymbol{y}\cdot\boldsymbol{u}_2}{\boldsymbol{u}_2\cdot\boldsymbol{u}_2}\boldsymbol{u}_2\\&=\frac{9}{30}\begin{bmatrix}2\\5\\-1\end{bmatrix}+\frac{3}{6}\begin{bmatrix}-2\\1\\1\end{bmatrix}=\frac{9}{30}\begin{bmatrix}2\\5\\-1\end{bmatrix}+\frac{15}{30}\begin{bmatrix}-2\\1\\1\end{bmatrix}=\begin{bmatrix}-2/5\\2\\1/5\end{bmatrix}\end{aligned}$$

且

$$\boldsymbol{y}-\hat{\boldsymbol{y}}=\begin{bmatrix}1\\2\\3\end{bmatrix}-\begin{bmatrix}-2/5\\2\\1/5\end{bmatrix}=\begin{bmatrix}7/5\\0\\14/5\end{bmatrix}$$

定理 8 保证$\boldsymbol{y}-\hat{\boldsymbol{y}}$属于$W^{\perp}$，为检验计算，一个好的方法是验证$\boldsymbol{y}-\hat{\boldsymbol{y}}$正交于$\boldsymbol{u}_1$和$\boldsymbol{u}_2$，因而正交于$W$，$\boldsymbol{y}$的期望分解式是

$$\boldsymbol{y}=\begin{bmatrix}1\\2\\3\end{bmatrix}=\begin{bmatrix}-2/5\\2\\1/5\end{bmatrix}+\begin{bmatrix}7/5\\0\\14/5\end{bmatrix}$$

■

正交投影的几何解释

当W是一维子空间时，公式（2）的$\mathrm{proj}_W\boldsymbol{y}$仅包含一项，这样，当$\dim W>1$，（2）式中的每一项都是自己，即向量$\boldsymbol{y}$在由$W$的一个基$\boldsymbol{u}$所生成的子空间上正交投影. 图 6-19 表示当$W$是

㊀ 我们可以假设W不是零子空间，否则$W^{\perp}=\mathbb{R}^n$，（1）即为$\boldsymbol{y}=\boldsymbol{0}+\boldsymbol{y}$. 下一节将说明$\mathbb{R}^n$中的任意非零子空间都有正交基.

$\mathbb{R}^3$ 中由 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 所生成子空间的情形，$\hat{\boldsymbol{y}}_1$ 和 $\hat{\boldsymbol{y}}_2$ 分别表示 $\boldsymbol{y}$ 在由 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 所生成直线上的投影，向量 $\boldsymbol{y}$ 的 W 上正交投影 $\hat{\boldsymbol{y}}$ 是 $\boldsymbol{y}$ 在两个相互正交的一维子空间分别投影之和，图 6-19 中的向量 $\hat{\boldsymbol{y}}$ 对应 6.2 节图 6-13 中的向量 $\boldsymbol{y}$，其原因是现在 $\hat{\boldsymbol{y}}$ 属于 W.

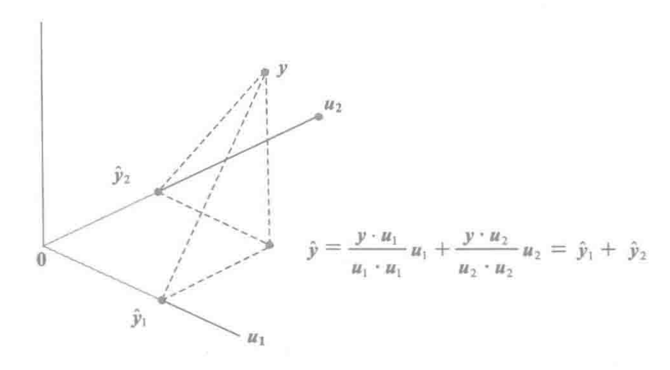


图 6-19 向量 $\boldsymbol{y}$ 的正交投影等于它在相互正交的一维子空间上的投影之和

正交投影的性质

如果 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是 W 的正交基，且假若 $\boldsymbol{y}$ 正好属于 W，那么 $\text{proj}_W\boldsymbol{y}$ 的公式和 6.2 节中定理 5 里 $\boldsymbol{y}$ 的表达式完全一致，这种情形下，$\text{proj}_W\boldsymbol{y}=\boldsymbol{y}$.

$$\text{如果 } \boldsymbol{y} \text{ 属于 } W=\text{Span}\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}\text{，那么 } \text{proj}_W\boldsymbol{y}=\boldsymbol{y}$$

这个结论可以从下列定理得出.

定理 9 （最佳逼近定理）

假设 W 是 $\mathbb{R}^n$ 空间中的一个子空间，$\boldsymbol{y}$ 是 $\mathbb{R}^n$ 中的任意向量，$\hat{\boldsymbol{y}}$ 是 $\boldsymbol{y}$ 在 W 上的正交投影，那么 $\hat{\boldsymbol{y}}$ 是 W 中最接近 $\boldsymbol{y}$ 的点，也就是指

$$\|\boldsymbol{y}-\hat{\boldsymbol{y}}\|<\|\boldsymbol{y}-\boldsymbol{v}\| \tag{3}$$

对所有属于 W 又异于 $\hat{\boldsymbol{y}}$ 的 $\boldsymbol{v}$ 成立.

定理 9 中的向量 $\hat{\boldsymbol{y}}$ 称为 **W 中元素对 $\boldsymbol{y}$ 的最佳逼近**，在后面章节中，我们会研究这样的问题，对给定的元素 $\boldsymbol{y}$，可以被某个给定子空间中的元素代替或“逼近”，用 $\|\boldsymbol{y}-\boldsymbol{v}\|$ 表示的从 $\boldsymbol{y}$ 到 $\boldsymbol{v}$ 的距离，可以认为是用 $\boldsymbol{v}$ 代替 $\boldsymbol{y}$ 的“误差”，定理 9 说明误差在 $\boldsymbol{v}=\hat{\boldsymbol{y}}$ 处取得最小值.

方程（3）同时给出一个结论，重新证明了 $\hat{\boldsymbol{y}}$ 的计算不依赖于特定正交基. 如果 W 的一个不同的正交基被用于构造 $\boldsymbol{y}$ 的正交投影，那么这个投影就是向量 $\boldsymbol{y}$ 在 W 中的最近点，记为 $\hat{\boldsymbol{y}}$.

证 取 W 中的 $\boldsymbol{v}$ 且不同于 $\hat{\boldsymbol{y}}$，见图 6-20，则 $\hat{\boldsymbol{y}}-\boldsymbol{v}$ 是 W 中的向量，由正交分解定理，$\boldsymbol{y}-\hat{\boldsymbol{y}}$ 正交于 W，特别地，$\boldsymbol{y}-\hat{\boldsymbol{y}}$ 正交于 $\hat{\boldsymbol{y}}-\boldsymbol{v}$，由于

$$\boldsymbol{y}-\boldsymbol{v}=(\boldsymbol{y}-\hat{\boldsymbol{y}})+(\hat{\boldsymbol{y}}-\boldsymbol{v})$$

由勾股定理可知

$$\|\boldsymbol{y}-\boldsymbol{v}\|^2=\|\boldsymbol{y}-\hat{\boldsymbol{y}}\|^2+\|\hat{\boldsymbol{y}}-\boldsymbol{v}\|^2$$

（见图 6-20 中右边的直角三角形.）因为 $\hat{\boldsymbol{y}}-\boldsymbol{v}\neq 0$，可知 $\|\hat{\boldsymbol{y}}-\boldsymbol{v}\|^2>0$，则不等式（3）立刻推出.

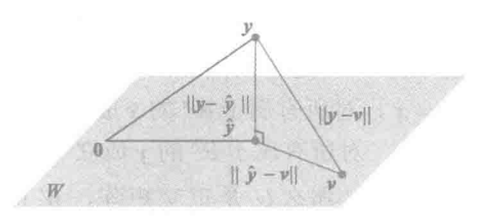


图 6-20　$\boldsymbol{y}$ 在 W 上的正交投影是 W 中离 $\boldsymbol{y}$ 最近的点 ■

例 3　如果 $\boldsymbol{u}_1=\begin{bmatrix}2\\5\\-1\end{bmatrix},\boldsymbol{u}_2=\begin{bmatrix}-2\\1\\1\end{bmatrix},\boldsymbol{y}=\begin{bmatrix}1\\2\\3\end{bmatrix}$ 和 $W=\text{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$，类似例 2，$W$ 中离 $\boldsymbol{y}$ 最接近的点是

$$\hat{\boldsymbol{y}}=\frac{\boldsymbol{y}\cdot\boldsymbol{u}_1}{\boldsymbol{u}_1\cdot\boldsymbol{u}_1}\boldsymbol{u}_1+\frac{\boldsymbol{y}\cdot\boldsymbol{u}_2}{\boldsymbol{u}_2\cdot\boldsymbol{u}_2}\boldsymbol{u}_2=\begin{bmatrix}-2/5\\2\\1/5\end{bmatrix}$$

■

例 4　$\mathbb{R}^n$ 中的一个点 $\boldsymbol{y}$ 到一个子空间 W 的距离，定义为从 $\boldsymbol{y}$ 到空间 W 中最近点的距离，求出 $\boldsymbol{y}$ 到 $W=\text{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$ 的距离，其中

$$\boldsymbol{y}=\begin{bmatrix}-1\\-5\\10\end{bmatrix},\quad \boldsymbol{u}_1=\begin{bmatrix}5\\-2\\1\end{bmatrix},\quad \boldsymbol{u}_2=\begin{bmatrix}1\\2\\-1\end{bmatrix}$$

解　由最佳逼近定理，从 $\boldsymbol{y}$ 到 W 的距离是 $\|\boldsymbol{y}-\hat{\boldsymbol{y}}\|$，且 $\hat{\boldsymbol{y}}=\text{proj}_W\boldsymbol{y}$，由于 $\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$ 是 W 的正交基，我们得到：

$$\hat{\boldsymbol{y}}=\frac{15}{30}\boldsymbol{u}_1+\frac{-21}{6}\boldsymbol{u}_2=\frac{1}{2}\begin{bmatrix}5\\-2\\1\end{bmatrix}-\frac{7}{2}\begin{bmatrix}1\\2\\-1\end{bmatrix}=\begin{bmatrix}-1\\-8\\4\end{bmatrix}$$

$$\boldsymbol{y}-\hat{\boldsymbol{y}}=\begin{bmatrix}-1\\-5\\10\end{bmatrix}-\begin{bmatrix}-1\\-8\\4\end{bmatrix}=\begin{bmatrix}0\\3\\6\end{bmatrix}$$

$$\|\boldsymbol{y}-\hat{\boldsymbol{y}}\|^2=3^2+6^2=45$$

从 $\boldsymbol{y}$ 到 W 的距离是 $\sqrt{45}=3\sqrt{5}$. ■

本节最后的定理说明，关于 $\text{proj}_W\boldsymbol{y}$ 的公式（2），当 W 的基是单位正交集时如何被简化.

定理 10　如果 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是 $\mathbb{R}^n$ 中子空间 W 的正交基，那么

$$\text{proj}_W\boldsymbol{y}=(\boldsymbol{y}\cdot\boldsymbol{u}_1)\boldsymbol{u}_1+(\boldsymbol{y}\cdot\boldsymbol{u}_2)\boldsymbol{u}_2+\cdots+(\boldsymbol{y}\cdot\boldsymbol{u}_p)\boldsymbol{u}_p \tag{4}$$

如果 $U=[\boldsymbol{u}_1\ \ \boldsymbol{u}_2\ \ \cdots\ \ \boldsymbol{u}_p]$，则

$$\text{proj}_W\boldsymbol{y}=UU^{\text{T}}\boldsymbol{y}\text{，对所有 }\boldsymbol{y}\in\mathbb{R}^n\text{ 成立} \tag{5}$$

证　公式（4）可由（2）式立刻推出，同样，（4）式说明 $\text{proj}_W\boldsymbol{y}$ 是 U 中列的线性组合，且对应权值分别为 $\boldsymbol{y}\cdot\boldsymbol{u}_1$，$\boldsymbol{y}\cdot\boldsymbol{u}_2$，$\cdots$，$\boldsymbol{y}\cdot\boldsymbol{u}_p$，同样可写成 $\boldsymbol{u}_1^{\text{T}}\boldsymbol{y}$，$\boldsymbol{u}_2^{\text{T}}\boldsymbol{y}$，$\cdots$，$\boldsymbol{u}_p^{\text{T}}\boldsymbol{y}$. 这表明它们可以记为

$U^{\mathrm{T}}\boldsymbol{y}$，从而证明了（5）式. ■

若 U 是 $n\times p$ 列单位正交的矩阵，W 是 U 的列空间，那么

$$U^{\mathrm{T}}U\boldsymbol{x}=I_p\boldsymbol{x}=\boldsymbol{x}\text{，对所有属于 }\mathbb{R}^p\text{ 的 }\boldsymbol{x}\text{ 成立} \qquad \text{定理 6}$$

$$UU^{\mathrm{T}}\boldsymbol{y}=\mathrm{proj}_W\boldsymbol{y}\text{，对所有属于 }\mathbb{R}^n\text{ 的 }\boldsymbol{y}\text{ 成立} \qquad \text{定理 10}$$

如果 U 是 $n\times n$ 方阵，且列单位正交，那么 U 是正交矩阵，W 的列空间是全部 $\mathbb{R}^n$ 空间，且 $UU^{\mathrm{T}}\boldsymbol{y}=I\boldsymbol{y}=\boldsymbol{y}$，对所有 $\boldsymbol{y}\in\mathbb{R}^n$ 成立.

尽管公式（4）在理论上十分重要，实际上，它常包含数字的平方根运算（在 $\boldsymbol{u}_i$ 的元素中），用手工运算时我们推荐公式（2）.

练习题

令 $\boldsymbol{u}_1=\begin{bmatrix}-7\\1\\4\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}-1\\1\\-2\end{bmatrix}$，$\boldsymbol{y}=\begin{bmatrix}-9\\1\\6\end{bmatrix}$，且 $W=\mathrm{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$，利用 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 正交计算 $\mathrm{proj}_W\boldsymbol{y}$.

习题 6.3

在习题 1 和习题 2 中，我们假设 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_4\}$ 是 $\mathbb{R}^4$ 中的正交基.

1. $\boldsymbol{u}_1=\begin{bmatrix}0\\1\\-4\\-1\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}3\\5\\1\\1\end{bmatrix}$，$\boldsymbol{u}_3=\begin{bmatrix}1\\0\\1\\-4\end{bmatrix}$，$\boldsymbol{u}_4=\begin{bmatrix}5\\-3\\-1\\1\end{bmatrix}$，$\boldsymbol{x}=\begin{bmatrix}10\\-8\\2\\0\end{bmatrix}$. 将 $\boldsymbol{x}$ 写成两个向量之和，一个属于 $\mathrm{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3\}$，另一个属于 $\mathrm{Span}\{\boldsymbol{u}_4\}$.

2. $\boldsymbol{u}_1=\begin{bmatrix}1\\2\\1\\1\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}-2\\1\\-1\\1\end{bmatrix}$，$\boldsymbol{u}_3=\begin{bmatrix}1\\1\\-2\\-1\end{bmatrix}$，$\boldsymbol{u}_4=\begin{bmatrix}-1\\1\\1\\-2\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}4\\5\\-3\\3\end{bmatrix}$. 将 $\boldsymbol{v}$ 写成两个向量之和，一个属于 $\mathrm{Span}\{\boldsymbol{u}_1\}$，另一个属于 $\mathrm{Span}\{\boldsymbol{u}_2,\boldsymbol{u}_3,\boldsymbol{u}_4\}$.

在习题 3~6 中，验证 $\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$ 是一个正交集，并找出 $\boldsymbol{y}$ 在由正交向量 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 所生成的子空间上的正交投影.

3. $\boldsymbol{y}=\begin{bmatrix}-1\\4\\3\end{bmatrix}$，$\boldsymbol{u}_1=\begin{bmatrix}1\\1\\0\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}-1\\1\\0\end{bmatrix}$

4. $\boldsymbol{y}=\begin{bmatrix}6\\3\\-2\end{bmatrix}$，$\boldsymbol{u}_1=\begin{bmatrix}3\\4\\0\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}-4\\3\\0\end{bmatrix}$

5. $\boldsymbol{y}=\begin{bmatrix}-1\\2\\6\end{bmatrix}$，$\boldsymbol{u}_1=\begin{bmatrix}3\\-1\\2\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}1\\-1\\-2\end{bmatrix}$

6. $\boldsymbol{y}=\begin{bmatrix}6\\4\\1\end{bmatrix}$，$\boldsymbol{u}_1=\begin{bmatrix}-4\\-1\\1\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}0\\1\\1\end{bmatrix}$

在习题 7~10 中，若 W 是由 $\boldsymbol{u}$ 组成的子空间，将 $\boldsymbol{y}$ 写成 W 中的向量与正交于 W 的向量之和.

7. $\boldsymbol{y}=\begin{bmatrix}1\\3\\5\end{bmatrix}$，$\boldsymbol{u}_1=\begin{bmatrix}1\\3\\-2\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}5\\1\\4\end{bmatrix}$

8. $\boldsymbol{y}=\begin{bmatrix}-1\\4\\3\end{bmatrix}$，$\boldsymbol{u}_1=\begin{bmatrix}1\\1\\1\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}-1\\3\\-2\end{bmatrix}$

9. $\boldsymbol{y}=\begin{bmatrix}4\\3\\3\\-1\end{bmatrix}$，$\boldsymbol{u}_1=\begin{bmatrix}1\\1\\0\\1\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}-1\\3\\1\\-2\end{bmatrix}$，$\boldsymbol{u}_3=\begin{bmatrix}-1\\0\\1\\1\end{bmatrix}$

10. $\boldsymbol{y}=\begin{bmatrix}3\\4\\5\\6\end{bmatrix}$，$\boldsymbol{u}_1=\begin{bmatrix}1\\1\\0\\-1\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}1\\0\\1\\1\end{bmatrix}$，$\boldsymbol{u}_3=\begin{bmatrix}0\\-1\\1\\-1\end{bmatrix}$

在习题 11 和习题 12 中，在 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 所生成的子空间 W 中，找出距离 $\boldsymbol{y}$ 最近的点.

11. $\boldsymbol{y}=\begin{bmatrix}3\\1\\5\\1\end{bmatrix}$，$\boldsymbol{v}_1=\begin{bmatrix}3\\1\\-1\\1\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}$

12. $\boldsymbol{y}=\begin{bmatrix}3\\-1\\1\\13\end{bmatrix}$，$\boldsymbol{v}_1=\begin{bmatrix}1\\-2\\-1\\2\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}-4\\1\\0\\3\end{bmatrix}$

在习题 13 和习题 14 中，在形如 $c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2$ 的向量中找出最接近 $\boldsymbol{z}$ 的向量.

13. $\boldsymbol{z}=\begin{bmatrix}3\\-7\\2\\3\end{bmatrix}$，$\boldsymbol{v}_1=\begin{bmatrix}2\\-1\\-3\\1\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}1\\1\\0\\-1\end{bmatrix}$

14. $\boldsymbol{z}=\begin{bmatrix}2\\4\\0\\-1\end{bmatrix}$，$\boldsymbol{v}_1=\begin{bmatrix}2\\0\\-1\\-3\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}5\\-2\\4\\2\end{bmatrix}$

15. 假设 $\boldsymbol{y}=\begin{bmatrix}5\\-9\\5\end{bmatrix}$，$\boldsymbol{u}_1=\begin{bmatrix}-3\\-5\\1\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}-3\\2\\1\end{bmatrix}$，找出向量 $\boldsymbol{y}$ 与 $\mathbb{R}^3$ 中 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 所生成平面的距离.

16. 若 $\boldsymbol{y}$，$\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 如习题 12 中所示，找出 $\boldsymbol{y}$ 到 $\mathbb{R}^4$ 中 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 所生成子空间的距离.

17. 设 $\boldsymbol{y}=\begin{bmatrix}4\\8\\1\end{bmatrix}$，$\boldsymbol{u}_1=\begin{bmatrix}2/3\\1/3\\2/3\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}-2/3\\2/3\\1/3\end{bmatrix}$，且 $W=\text{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$.

a. 若 $U=[\boldsymbol{u}_1\ \ \boldsymbol{u}_2]$，计算 $U^{\mathrm{T}}U$ 和 UU^{T}.

b. 计算 $\text{proj}_W\boldsymbol{y}$ 和 $(UU^{\mathrm{T}})\boldsymbol{y}$.

18. 设 $\boldsymbol{y}=\begin{bmatrix}7\\9\end{bmatrix}$，$\boldsymbol{u}_1=\begin{bmatrix}1/\sqrt{10}\\-3/\sqrt{10}\end{bmatrix}$ 和 $W=\text{Span}\{\boldsymbol{u}_1\}$.

a. 若 U 是 2×1 矩阵，其列向量为 $\boldsymbol{u}_1$，计算 $U^{\mathrm{T}}U$ 和 UU^{T}.

b. 计算 $\text{proj}_W\boldsymbol{y}$ 和 $(UU^{\mathrm{T}})\boldsymbol{y}$.

19. 假设 $\boldsymbol{u}_1=\begin{bmatrix}1\\1\\-2\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}5\\-1\\2\end{bmatrix}$ 和 $\boldsymbol{u}_3=\begin{bmatrix}0\\0\\1\end{bmatrix}$，注意 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 正交，但 $\boldsymbol{u}_3$ 和 $\boldsymbol{u}_1$ 或 $\boldsymbol{u}_2$ 不正交，可以证明 $\boldsymbol{u}_3$ 不属于 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 所生成的子空间，利用这个事实，构造 $\mathbb{R}^3$ 的非零向量 $\boldsymbol{v}$ 与 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 正交.

20. 设 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 如习题 19 所示，且 $\boldsymbol{u}_4=\begin{bmatrix}0\\1\\0\end{bmatrix}$，可以验证 $\boldsymbol{u}_4$ 不属于 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 所生成的子空间 W，利用这个结论构造 $\mathbb{R}^3$ 的非零向量 $\boldsymbol{v}$ 与 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 正交.

在习题 21~22 中，所有向量和子空间均属于 $\mathbb{R}^n$，判断下列结论的正误，验证每个结论.

21. a. 如果 $\boldsymbol{z}$ 与 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 正交，且 $W=\text{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$，那么 $\boldsymbol{z}$ 一定属于 $W^{\perp}$.

b. 对任一向量 $\boldsymbol{y}$ 和任一子空间 W，向量 $\boldsymbol{y}-\text{proj}_W\boldsymbol{y}$ 正交于 W.

c. $\boldsymbol{y}$ 在子空间 W 的正交投影 $\hat{\boldsymbol{y}}$ 的计算，有时可能依赖于 W 正交基的选取.

d. 如果 $\boldsymbol{y}$ 属于子空间 W，那么 $\boldsymbol{y}$ 在 W 的正交投影是 $\boldsymbol{y}$ 本身.

e. 如果 $n\times p$ 矩阵 U 的列是单位正交的，那么 $UU^{\mathrm{T}}\boldsymbol{y}$ 是 $\boldsymbol{y}$ 在 U 的列空间上的正交投影.

22. a. 如果 W 是 $\mathbb{R}^n$ 的子空间，且 $\boldsymbol{v}$ 属于 W 和 $W^{\perp}$，那么 $\boldsymbol{v}$ 一定是零向量.

b. 在正交分解定理中，公式（2）中 $\hat{\boldsymbol{y}}$ 的每一项在子空间 W 上的正交投影是它本身.

c. 如果 $\boldsymbol{y}=\boldsymbol{z}_1+\boldsymbol{z}_2$，其中 $\boldsymbol{z}_1$ 属于子空间 W 且 $\boldsymbol{z}_2$ 属于子空间 $W^{\perp}$，那么 $\boldsymbol{z}_1$ 一定是 $\boldsymbol{y}$ 在 W 上的正交投影.

d. 子空间 W 中的元素对向量 $\boldsymbol{y}$ 的最佳逼近是向量 $\boldsymbol{y}-\text{proj}_W\boldsymbol{y}$.

e. 如果一个 $n\times p$ 矩阵 U 有单位正交列，那么 $UU^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{x}$，对所有 $\boldsymbol{x}\in\mathbb{R}^n$ 成立.

23. 若 A 是 $m\times n$ 矩阵，证明 $\mathbb{R}^n$ 中任一向量 $\boldsymbol{x}$ 可以表示写成 $\boldsymbol{x}=\boldsymbol{p}+\boldsymbol{u}$，此处，$\boldsymbol{p}$ 属于 Row A 且 $\boldsymbol{u}$

属于 Nul A .同样可以证明：如果方程 $Ax=b$ 是相容的，那么存在唯一向量 p 属于 Row A，使得 $Ap=b$.

24. 假若 W 是 $\mathbb{R}^n$ 的一个子空间且 $\{w_1,\cdots,w_p\}$ 是 W 的一个正交基，$\{v_1,\cdots,v_q\}$ 是 $W^\perp$ 的正交基.
    a. 说明为什么 $\{w_1,\cdots,w_p,v_1,\cdots,v_q\}$ 是正交集.
    b. 说明为什么（a）中的集合可以生成 $\mathbb{R}^n$.
    c. 证明 $\dim W+\dim W^\perp=n$.
25. [M] 若 U 是 6.2 节中习题 36 题的 8×4 矩阵，在 Col U 中找出最接近 $y=(1,1,1,1,1,1,1,1)$ 的点，写出你解决该问题所使用的键击内容或命令.
26. [M] 假设 U 是习题 25 中的矩阵，求出从 $b=(1,1,1,1,-1,-1,-1,-1)$ 到 Col U 的距离.

练习题答案

计算

$$\mathrm{proj}_W y=\frac{y\cdot u_1}{u_1\cdot u_1}u_1+\frac{y\cdot u_2}{u_2\cdot u_2}u_2=\frac{88}{66}u_1+\frac{-2}{6}u_2$$
$$=\frac{4}{3}\begin{bmatrix}-7\\1\\4\end{bmatrix}-\frac{1}{3}\begin{bmatrix}-1\\1\\-2\end{bmatrix}=\begin{bmatrix}-9\\1\\6\end{bmatrix}=y$$

在这种情形下，y 恰好是 u_1 和 u_2 的线性组合，从而 y 属于 W，即 W 中最接近 y 的点是 y 本身.

6.4 格拉姆-施密特方法

格拉姆-施密特方法是对 $\mathbb{R}^n$ 中任何非零子空间，构造正交基或标准正交基的简单算法，本节前面两个例题顺便给出计算步骤.

例 1 假设 $W=\mathrm{Span}\{x_1,x_2\}$，其中 $x_1=\begin{bmatrix}3\\6\\0\end{bmatrix}$，$x_2=\begin{bmatrix}1\\2\\2\end{bmatrix}$. 构造 $W=\mathrm{Span}\{x_1,x_2\}$ 的一个正交基.

解 子空间 W 如图 6-21 所示，沿着 x_1，x_2 和 x_2 在 x_1 上的投影 p，与 x_1 正交的 x_2 的分量是 x_2-p，它属于 W，其原因是它由 x_2 和 x_1 的倍数相加得到，取 $v_1=x_1$，可得

$$v_2=x_2-p=x_2-\frac{x_2\cdot x_1}{x_1\cdot x_1}\cdot x_1=\begin{bmatrix}1\\2\\2\end{bmatrix}-\frac{15}{45}\begin{bmatrix}3\\6\\0\end{bmatrix}=\begin{bmatrix}0\\0\\2\end{bmatrix}$$

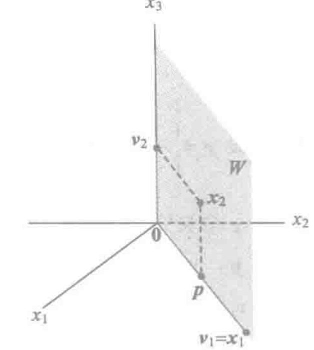


图 6-21　正交基 $\{v_1,v_2\}$ 的构造

那么 $\{v_1,v_2\}$ 是空间 W 的非零向量构成的正交集，由于 $\dim W=2$，可知 $\{v_1,v_2\}$ 构成 W 的一个基.■

下面的例子详细说明格拉姆-施密特方法，请仔细学习.

例 2 设 $x_1=\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$，$x_2=\begin{bmatrix}0\\1\\1\\1\end{bmatrix}$，$x_3=\begin{bmatrix}0\\0\\1\\1\end{bmatrix}$. 那么 $\{x_1,x_2,x_3\}$ 显然线性无关，且构成 $\mathbb{R}^4$ 中子空间

W 的一个基，试构造 W 的一个正交基.

解 **步骤** 1. 取 $\boldsymbol{v}_1 = \boldsymbol{x}_1$ 和 $W_1 = \text{Span}\{\boldsymbol{x}_1\} = \text{Span}\{\boldsymbol{v}_1\}$.

步骤 2. 取 $\boldsymbol{v}_2$ 是 $\boldsymbol{x}_2$ 减去它在子空间 W_1 上的投影所得到的向量，即

$$\begin{aligned}\boldsymbol{v}_2 &= \boldsymbol{x}_2 - \text{proj}_{W_1}\boldsymbol{x}_2 \\ &= \boldsymbol{x}_2 - \frac{\boldsymbol{x}_2 \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 \text{（因为 } \boldsymbol{v}_1 = \boldsymbol{x}_1\text{）} \\ &= \begin{bmatrix}0\\1\\1\\1\end{bmatrix} - \frac{3}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} = \begin{bmatrix}-3/4\\1/4\\1/4\\1/4\end{bmatrix}\end{aligned}$$

像例 1 一样，$\boldsymbol{v}_2$ 是 $\boldsymbol{x}_2$ 正交于 $\boldsymbol{x}_1$ 的分量，且 $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ 是由 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 所生成空间 W_2 的一个正交基.

步骤 2′（可选的）. 如果合适，度量 $\boldsymbol{v}_2$ 可简化后面的计算. 由于 $\boldsymbol{v}_2$ 具有分数分量，很方便用因子 4 重新度量向量 $\boldsymbol{v}_2$，用下面两个正交基代替 $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$.

$$\boldsymbol{v}_1 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \boldsymbol{v}_2' = \begin{bmatrix}-3\\1\\1\\1\end{bmatrix}$$

步骤 3. 取 $\boldsymbol{v}_3$ 是 $\boldsymbol{x}_3$ 减去它在子空间 W_2 上的投影所产生的向量，先利用正交基 $\{\boldsymbol{v}_1, \boldsymbol{v}_2'\}$ 计算 $\boldsymbol{x}_3$ 在 W_2 上的投影.

$$\begin{aligned}\text{proj}_{W_2}\boldsymbol{x}_3 &= \underbrace{\frac{\boldsymbol{x}_3 \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1}_{\boldsymbol{x}_3\text{ 在 }\boldsymbol{v}_1\text{ 上的投影}} + \underbrace{\frac{\boldsymbol{x}_3 \cdot \boldsymbol{v}'_2}{\boldsymbol{v}_2' \cdot \boldsymbol{v}_2'}\boldsymbol{v}_2'}_{\boldsymbol{x}_3\text{ 在 }\boldsymbol{v}_2'\text{ 上的投影}} \\ &= \frac{2}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} + \frac{2}{12}\begin{bmatrix}-3\\1\\1\\1\end{bmatrix} = \begin{bmatrix}0\\2/3\\2/3\\2/3\end{bmatrix}\end{aligned}$$

那么 $\boldsymbol{v}_3$ 是 $\boldsymbol{x}_3$ 在 W_2 上正交分量，具体为：

$$\boldsymbol{v}_3 = \boldsymbol{x}_3 - \text{proj}_{W_2}\boldsymbol{x}_3 = \begin{bmatrix}0\\0\\1\\1\end{bmatrix} - \begin{bmatrix}0\\2/3\\2/3\\2/3\end{bmatrix} = \begin{bmatrix}0\\-2/3\\1/3\\1/3\end{bmatrix}$$

从图 6-22 可以看到构造的图解，观察可得 $\boldsymbol{v}_3$ 属于 W，原因是 $\boldsymbol{x}_3$ 和 $\text{proj}_{W_2}\boldsymbol{x}_3$ 都属于 W. 从而 $\{\boldsymbol{v}_1, \boldsymbol{v}_2', \boldsymbol{v}_3\}$ 是 W 中非零向量构成的正交集（因而是线性无关集），由于 W 是三维空间，它的基又包含三个向量，所以由 4.5 节的基定理可知，$\{\boldsymbol{v}_1, \boldsymbol{v}_2', \boldsymbol{v}_3\}$ 是 W 的正交基. ■

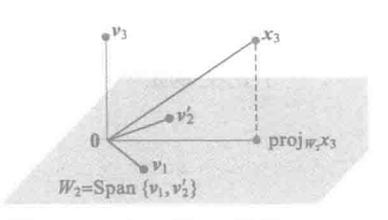


图 6-22 从 $\boldsymbol{x}_3$ 和 W_2 构造出 $\boldsymbol{v}_3$

下面定理的证明表示这种方法确实有效，重新度量的步骤没有涉及，原因是度量只对手工计算作简化.

定理 11 （格拉姆-施密特方法）

对 $\mathbb{R}^n$ 中子空间的一个基 $\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_p\}$，定义

$$\boldsymbol{v}_1=\boldsymbol{x}_1$$

$$\boldsymbol{v}_2=\boldsymbol{x}_2-\frac{\boldsymbol{x}_2\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1}\boldsymbol{v}_1$$

$$\boldsymbol{v}_3=\boldsymbol{x}_3-\frac{\boldsymbol{x}_3\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1}\boldsymbol{v}_1-\frac{\boldsymbol{x}_3\cdot\boldsymbol{v}_2}{\boldsymbol{v}_2\cdot\boldsymbol{v}_2}\boldsymbol{v}_2$$

$$\vdots$$

$$\boldsymbol{v}_p=\boldsymbol{x}_p-\frac{\boldsymbol{x}_p\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1}\boldsymbol{v}_1-\frac{\boldsymbol{x}_p\cdot\boldsymbol{v}_2}{\boldsymbol{v}_2\cdot\boldsymbol{v}_2}\boldsymbol{v}_2-\cdots-\frac{\boldsymbol{x}_p\cdot\boldsymbol{v}_{p-1}}{\boldsymbol{v}_{p-1}\cdot\boldsymbol{v}_{p-1}}\boldsymbol{v}_{p-1}$$

那么 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 是 W 的一个正交基，此外

$$\text{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k\}=\text{Span}\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k\}，\text{其中 } 1\leqslant k\leqslant p \tag{1}$$

证 对 $1\leqslant k\leqslant p$，取 $W_k=\text{Span}\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k\}$.让 $\boldsymbol{v}_1=\boldsymbol{x}_1$，则有 $\text{Span}\{\boldsymbol{v}_1\}=\text{Span}\{\boldsymbol{x}_1\}$. 假若对 $k<p$，我们已经构造 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k$，使得 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k\}$ 是 W_k 的一个正交基，定义

$$\boldsymbol{v}_{k+1}=\boldsymbol{x}_{k+1}-\text{proj}_{W_k}\boldsymbol{x}_{k+1} \tag{2}$$

由正交分解定理知，$\boldsymbol{v}_{k+1}$ 正交于 W_k. 注意向量 $\text{proj}_{W_k}\boldsymbol{x}_{k+1}$ 属于 W_k，因而属于 W_{k+1}，又由于 $\boldsymbol{x}_{k+1}$ 属于 W_{k+1}，从而推出 $\boldsymbol{v}_{k+1}$ 也属于 W_{k+1}（因为 W_{k+1} 是子空间，且减法封闭）. 更进一步，由于 $\boldsymbol{x}_{k+1}$ 不属于 $W_k=\text{Span}\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k\}$ 可得 $\boldsymbol{v}_{k+1}\neq\boldsymbol{0}$.因而 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_{k+1}\}$ 是 $(k+1)$ 维空间 W_{k+1} 中非零向量形成的正交集，由 4.5 节的基定理可知，这个集合是 W_{k+1} 的正交基，从而 $W_{k+1}=\text{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_{k+1}\}$，当 $k+1=p$ 时成立，归纳证明过程结束. ■

定理 11 说明任何 $\mathbb{R}^n$ 中的非零子空间 W 有一个正交集，因为普通的基 $\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_p\}$ 总是存在的（由 4.5 节中定理 11），而格拉姆-施密特方法仅依赖于在有正交基的子空间 W 上的正交投影的存在性.

标准正交基

一个标准正交基很容易从一个正交基 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 得到：只需单位化（度量）所有 $\boldsymbol{v}_k$. 当问题用手计算时，它比一开始找到 $\boldsymbol{v}_k$ 就将其单位化更容易（因为可避免不必要的平方根运算）.

例 3 在例 1，我们构造正交集

$$\boldsymbol{v}_1=\begin{bmatrix}3\\6\\0\end{bmatrix},\ \boldsymbol{v}_2=\begin{bmatrix}0\\0\\2\end{bmatrix}$$

一个正交基是

$$\boldsymbol{u}_1=\frac{1}{\|\boldsymbol{v}_1\|}\boldsymbol{v}_1=\frac{1}{\sqrt{45}}\begin{bmatrix}3\\6\\0\end{bmatrix}=\begin{bmatrix}1/\sqrt{5}\\2/\sqrt{5}\\0\end{bmatrix}$$

$$u_2 = \frac{1}{\|v_2\|} v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$ ■

矩阵的 QR 分解

如果 $m \times n$ 矩阵 A 的列 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ 线性无关，那么应用格拉姆-施密特方法（包含单位化）于 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ 等同于按下面定理描述的方法分解 A，这种分解方法广泛应用在各种计算机算法中，如解方程（6.5 节讨论）和求特征值（5.2 节的习题提到）.

定理 12 （QR 分解）

如果 $m \times n$ 矩阵 A 的列线性无关，那么 A 可以分解为 $A = QR$，其中 Q 是一个 $m \times n$ 矩阵，其列形成 $\mathrm{Col}\, A$ 的一个标准正交基，R 是一个 $n \times n$ 上三角可逆矩阵且在对角线上的元素为正数.

证 A 的列向量形成 $\mathrm{Col}\, A$ 的一个基 $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$，构造 $W = \mathrm{Col}\, A$ 的一个标准正交基 $\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n\}$，且具有定理 11 中的性质（1），这个基的构造可由格拉姆-施密特方法或其他方法完成，取

$$Q = [\boldsymbol{u}_1 \ \ \boldsymbol{u}_2 \ \ \cdots \ \ \boldsymbol{u}_n]$$

对 $k = 1, \cdots, n$，$\boldsymbol{x}_k$ 属于 $\mathrm{Span}\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k\} = \mathrm{Span}\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_k\}$.所以存在常数 $r_{1k}, \cdots, r_{kk}$，使得

$$\boldsymbol{x}_k = r_{1k}\boldsymbol{u}_1 + \cdots + r_{kk}\boldsymbol{u}_k + 0 \cdot \boldsymbol{u}_{k+1} + \cdots + 0 \cdot \boldsymbol{u}_n$$

我们可以假设 $r_{kk} \geqslant 0$（如果 $r_{kk} < 0$，对 r_{kk} 和 $\boldsymbol{u}_k$ 都乘-1），这表明 $\boldsymbol{x}_k$ 是 Q 中列的线性组合，且系数是下面向量的分量：

$$\boldsymbol{r}_k = \begin{bmatrix} r_{1k} \\ \vdots \\ r_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

即 $\boldsymbol{x}_k = Q\boldsymbol{r}_k$，其中 $k = 1, \cdots, n$，取 $R = [\boldsymbol{r}_1 \cdots \boldsymbol{r}_n]$. 那么

$$A = [\boldsymbol{x}_1 \cdots \boldsymbol{x}_n] = [Q\boldsymbol{r}_1 \cdots Q\boldsymbol{r}_n] = QR$$

R 可逆的结论，可从下面 $\mathrm{Col}\, A$ 线性无关的事实立刻得到（习题 19），因为 R 是上三角形矩阵，它的非负对角元素必为正值. ■

例 4 求 $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ 的一个 QR 分解.

解 A 的列向量为例 2 中的 $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$，$\mathrm{Col}\, A = \mathrm{Span}\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\}$ 的一个正交基在例 2 中得到：

$$v_1=\begin{bmatrix}1\\1\\1\\1\end{bmatrix},\quad v_2'=\begin{bmatrix}-3\\1\\1\\1\end{bmatrix},\quad v_3=\begin{bmatrix}0\\-2/3\\1/3\\1/3\end{bmatrix}$$

重新度量v_3，取$v_3'=3v_3$，那么将v_1,v_2,v_3三个向量单位化得到u_1,u_2,u_3，且用这些向量组成Q的列.

$$Q=\begin{bmatrix}1/2 & -3/\sqrt{12} & 0\\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6}\\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6}\\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6}\end{bmatrix}$$

由Q的构造可知，Q的前k列是$\mathrm{Span}\{x_1,\cdots,x_k\}$的一个标准正交基. 从定理12的证明可知，对某些$R$有$A=QR$. 为找到$R$，注意到$Q^{\mathrm{T}}Q=I$以及$Q$的列是单位正交向量，所以

$$Q^{\mathrm{T}}A=Q^{\mathrm{T}}(QR)=IR=R.$$

且

$$R=\begin{bmatrix}1/2 & 1/2 & 1/2 & 1/2\\ -3/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12}\\ 0 & -2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6}\end{bmatrix}\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\\1&1&1\end{bmatrix}$$

$$=\begin{bmatrix}2 & 3/2 & 1\\ 0 & 3/\sqrt{12} & 2/\sqrt{12}\\ 0 & 0 & 2/\sqrt{6}\end{bmatrix}$$

■

数值计算的注解

1. 当格拉姆-施密特方法用计算机实现时，在u_k的计算中，每个向量由四舍五入引起的误差逐渐增大，对较大但不相等的j和k，标量乘法$u_j^{\mathrm{T}}u_k$也许不充分接近零，通过重新安排计算的阶，这类正交性的损失可以大量减少.[㊀] 然而，QR分解常指这类标准化的格拉姆-施密特方法，因为它会产生更精确的标准正交基，即使分解需要双倍的算术计算.

2. 为了得到一个矩阵A的QR分解，计算机程序常常对A左乘一系列正交矩阵使结果变成一个上三角形矩阵，这个构造过程类似于A左乘一系列初等矩阵，最后得到A的LU分解.

练习题

设$W=\mathrm{Span}\{x_1,x_2\}$，此处$x_1=\begin{bmatrix}1\\1\\1\end{bmatrix}$和$x_2=\begin{bmatrix}1/3\\1/3\\-2/3\end{bmatrix}$，构造$W$的一个标准正交基.

㊀ 参见 *Fundamentals of Matrix Computations*, David S. Watkins (New York: John Wiley & Sons, 1991), pp. 167-180.

习题 6.4

在习题 1~6 中，给出的集合是子空间 W 的一个基，利用格拉姆-施密特方法构造 W 的正交基.

1. $\begin{bmatrix}3\\0\\-1\end{bmatrix},\begin{bmatrix}8\\5\\-6\end{bmatrix}$

2. $\begin{bmatrix}0\\4\\2\end{bmatrix},\begin{bmatrix}5\\6\\-7\end{bmatrix}$

3. $\begin{bmatrix}2\\-5\\1\end{bmatrix},\begin{bmatrix}4\\-1\\2\end{bmatrix}$

4. $\begin{bmatrix}3\\-4\\5\end{bmatrix},\begin{bmatrix}-3\\14\\-7\end{bmatrix}$

5. $\begin{bmatrix}1\\-4\\0\\1\end{bmatrix},\begin{bmatrix}7\\-7\\-4\\1\end{bmatrix}$

6. $\begin{bmatrix}3\\-1\\2\\-1\end{bmatrix},\begin{bmatrix}-5\\9\\-9\\3\end{bmatrix}$

7. 求习题 3 中向量所生成子空间的一个标准正交基.

8. 求习题 4 中向量所生成子空间的一个标准正交基.

在习题 9~12 中，求每个矩阵列空间的正交基.

9. $\begin{bmatrix}3&-5&1\\1&1&1\\-1&5&-2\\3&-7&8\end{bmatrix}$

10. $\begin{bmatrix}-1&6&6\\3&-8&3\\1&-2&6\\1&-4&-3\end{bmatrix}$

11. $\begin{bmatrix}1&2&5\\-1&1&-4\\-1&4&-3\\1&-4&7\\1&2&1\end{bmatrix}$

12. $\begin{bmatrix}1&3&5\\-1&-3&1\\0&2&3\\1&5&2\\1&5&8\end{bmatrix}$

在习题 13~14 中，矩阵 Q 的列是格拉姆-施密特方法应用于矩阵 A 的列而得到的，求上三角形矩阵 R，使得 $A=QR$，并检查你的结果.

13. $A=\begin{bmatrix}5&9\\1&7\\-3&-5\\1&5\end{bmatrix}$，$Q=\begin{bmatrix}5/6&-1/6\\1/6&5/6\\-3/6&1/6\\1/6&3/6\end{bmatrix}$

14. $A=\begin{bmatrix}-2&3\\5&7\\2&-2\\4&6\end{bmatrix}$，$Q=\begin{bmatrix}-2/7&5/7\\5/7&2/7\\2/7&-4/7\\4/7&2/7\end{bmatrix}$

15. 求习题 11 中矩阵的 QR 分解.

16. 求习题 12 中矩阵的 QR 分解.

习题 17~18 中的所有向量和子空间都属于 $\mathbb{R}^n$，判断每一个结论是正确还是错误，验证你的结论.

17. a. 如果 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 是 W 的正交基，那么用因子 c 去乘 $\boldsymbol{v}_3$ 可得一个新的正交基 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,c\boldsymbol{v}_3\}$.

    b. 格拉姆-施密特方法将一个线性无关的集合 $\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_p\}$ 转化为一个正交集 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$，且具有对每一个 k，向量 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k$ 生成的子空间与向量 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k$ 生成的子空间一致.

    c. 如果 $A=QR$，这里 Q 具有单位正交列，那么 $R=Q^{\mathrm{T}}A$.

18. a. 如果 $W=\mathrm{Span}\{\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3\}$，且 $\{\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3\}$ 线性无关，如果 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 是 W 的一个正交集，那么 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 是 W 的一个基.

    b. 如果 $\boldsymbol{x}$ 不属于子空间 W，那么 $\boldsymbol{x}-\mathrm{proj}_W\boldsymbol{x}$ 不是零向量.

    c. 在一个 QR 分解中，例如 $A=QR$，Q 的列构成 A 的列子空间的标准正交基.

19. 假设 $A=QR$，此处 Q 是 $m\times n$ 矩阵，R 是 $n\times n$ 矩阵. 证明：如果 A 的列线性无关，那么 R 一定可逆（提示：研究方程 $R\boldsymbol{x}=\boldsymbol{0}$ 且利用 $A=QR$ 的事实）.

20. 假设 $A=QR$，此处 R 是一个可逆矩阵，证明 A 和 Q 具有相同的列空间.（提示：对给定属于列 A 的 $\boldsymbol{y}$，存在 $\boldsymbol{x}$ 使得 $\boldsymbol{y}=Q\boldsymbol{x}$. 此外，对给定 $\boldsymbol{y}$ 属于 Col Q，存在 $\boldsymbol{x}$ 使得 $\boldsymbol{y}=A\boldsymbol{x}$.）

21. 如定理 12 中 $A=QR$，说明如何找到一个正交 $m\times m$（方）矩阵 Q_1 和一个可逆 $n\times n$ 上三角形矩阵 R，使得

$$A=Q_1\begin{bmatrix}R\\0\end{bmatrix}$$

当 rank $A=n$ 时，MATLAB 中的 qr 命令给出一个完全 QR 分解.

22. 设 $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p$ 是 $\mathbb{R}^n$ 上子空间 W 的一个正交基，$T:\mathbb{R}^n\to\mathbb{R}^n$ 定义为 $T(\boldsymbol{x})=\text{proj}_W\boldsymbol{x}$，证明 T 是一个线性变换.
23. 设 $A=QR$ 是 $m\times n$ 矩阵 A（含有线性无关列）的一个 QR 分解. 把 A 化分成 $[A_1\ \ A_2]$，这里 A_1 有 p 列. 说明如何获得 A_1 的一个 QR 分解，并解释为什么这样的分解具有如此性质.
24. [M]像例 2 那样利用格拉姆-施密特方法，构造下面矩阵 A 的列空间的正交基.
$$A=\begin{bmatrix}-10 & 13 & 7 & -11\\ 2 & 1 & -5 & 3\\ -6 & 3 & 13 & -3\\ 16 & -16 & -2 & 5\\ 2 & 1 & -5 & -7\end{bmatrix}$$
25. [M]利用本节中方法给出习题 24 中矩阵 A 的一个 QR 分解.
26. [M]对矩阵算法，格拉姆-施密特方法比单位正交向量更有效，从定理 11 中的 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_p$ 开始，取 $A=[\boldsymbol{x}_1\ \cdots\ \boldsymbol{x}_p]$，若 $n\times k$ 矩阵 Q_1 的列，构成矩阵 A 前 k 列所生成的子空间 W_k 的一个标准正交基，那么对于 $\mathbb{R}^n$ 中的向量 $\boldsymbol{x}$，$QQ^{\mathrm{T}}\boldsymbol{x}$ 是 $\boldsymbol{x}$ 在 W_k 上的正交投影（6.3 节定理 10）. 如果 $\boldsymbol{x}_{k+1}$ 是 A 的下一列，定理 11 中方程（2）的证明变成
$$\boldsymbol{v}_{k+1}=\boldsymbol{x}_{k+1}-Q(Q^{\mathrm{T}}\boldsymbol{x}_{k+1})$$
（上面的括号可以减少算术运算.）取 $\boldsymbol{u}_{k+1}=\boldsymbol{v}_{k+1}/\|\boldsymbol{v}_{k+1}\|$，下一个步骤中的新 Q 是 $[Q\ \ \boldsymbol{u}_{k+1}]$，利用这个步骤，计算习题 24 中矩阵的 QR 分解，写出你所用的键盘输入或命令.

练习题答案

设 $\boldsymbol{v}_1=\boldsymbol{x}_1=\begin{bmatrix}1\\1\\1\end{bmatrix}$ 和 $\boldsymbol{v}_2=\boldsymbol{x}_2-\dfrac{\boldsymbol{x}_2\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1}\boldsymbol{v}_1=\boldsymbol{x}_2-0\cdot\boldsymbol{v}_1=\boldsymbol{x}_2$，这样 $\{\boldsymbol{x}_1,\boldsymbol{x}_2\}$ 是正交向量，接下来需要向量的单位化. 设
$$\boldsymbol{u}_1=\frac{1}{\|\boldsymbol{v}_1\|}\boldsymbol{v}_1=\frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{bmatrix}$$
用 $\boldsymbol{v}_2'=3\boldsymbol{v}_2$ 代替第 2 个向量，
$$\boldsymbol{u}_2=\frac{1}{\|\boldsymbol{v}_2'\|}\boldsymbol{v}_2'=\frac{1}{\sqrt{1^2+1^2+(-2)^2}}\begin{bmatrix}1\\1\\-2\end{bmatrix}=\begin{bmatrix}1/\sqrt{6}\\1/\sqrt{6}\\-2/\sqrt{6}\end{bmatrix}$$
这样 $\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$ 就是 W 的标准正交基.

6.5 最小二乘问题

本章前言中的例子，描述一个解不存在的巨型方程组 $A\boldsymbol{x}=\boldsymbol{b}$，实际应用中常出现这类不相容问题，尽管不会出现如何巨大的系数矩阵. 当方程组的解不存在但又需要求解时，最好的方法是寻找 $\boldsymbol{x}$，使得 $A\boldsymbol{x}$ 尽可能接近 $\boldsymbol{b}$.

考虑 $A\boldsymbol{x}$ 作为 $\boldsymbol{b}$ 的一个近似，从 $\boldsymbol{b}$ 到 $A\boldsymbol{x}$ 的距离越小，$\|\boldsymbol{b}-A\boldsymbol{x}\|$ 近似程度越好. **一般的最小二乘问题**就是找出使 $\|\boldsymbol{b}-A\boldsymbol{x}\|$ 尽量小的 $\boldsymbol{x}$，术语“最小二乘”来源于这样的事实，$\|\boldsymbol{b}-A\boldsymbol{x}\|$ 是平方和的平方根.

定义 如果 $m\times n$ 矩阵 A 和向量 $\boldsymbol{b}$ 属于 $\mathbb{R}^m$，$A\boldsymbol{x}=\boldsymbol{b}$ 的**最小二乘解**是 $\mathbb{R}^n$ 中的 $\hat{\boldsymbol{x}}$，使得

$$\|\boldsymbol{b}-A\hat{\boldsymbol{x}}\|\leqslant\|\boldsymbol{b}-A\boldsymbol{x}\|$$

对所有 $\boldsymbol{x}\in\mathbb{R}^n$ 成立.

最小二乘问题最重要的特征是无论怎么选取 $\boldsymbol{x}$，向量 $A\boldsymbol{x}$ 必然属于列空间 Col A. 就是我们寻求 $\boldsymbol{x}$，使得 $A\boldsymbol{x}$ 是 Col A 中最接近 $\boldsymbol{b}$ 的向量，见图 6-23（当然，如果 $\boldsymbol{b}$ 恰好在 Col A 中，那么 $\boldsymbol{b}$ 就等于 $A\boldsymbol{x}$ 且这样的 $\boldsymbol{x}$ 也是一个“最小二乘解”）.

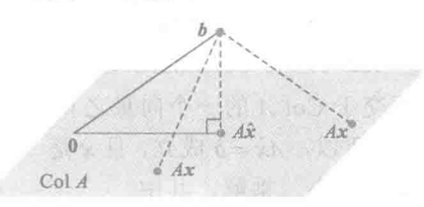


图 6-23 对 $\boldsymbol{x}$，$\boldsymbol{b}$ 与 $A\hat{\boldsymbol{x}}$ 的距离小于与其他 $A\boldsymbol{x}$ 的距离

一般最小二乘问题的解

对上面给定的 A 和 $\boldsymbol{b}$，应用 6.3 节的最佳逼近定理于 Col A 空间，取

$$\hat{\boldsymbol{b}}=\operatorname{proj}_{\operatorname{Col} A}\boldsymbol{b}$$

由于 $\hat{\boldsymbol{b}}$ 属于 A 的列空间，方程 $A\boldsymbol{x}=\hat{\boldsymbol{b}}$ 是相容的且存在一个属于 $\mathbb{R}^n$ 的 $\hat{\boldsymbol{x}}$ 使得

$$A\hat{\boldsymbol{x}}=\hat{\boldsymbol{b}} \tag{1}$$

由于 $\hat{\boldsymbol{b}}$ 是 Col A 中最接近 $\boldsymbol{b}$ 的点，一个向量 $\hat{\boldsymbol{x}}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的一个最小二乘解的充分必要条件是 $\hat{\boldsymbol{x}}$ 满足（1）. 这个属于 $\mathbb{R}^n$ 的 $\hat{\boldsymbol{x}}$ 是一系列由列 A 构造的 $\hat{\boldsymbol{b}}$ 的权值，见图 6-24（如果方程（1）有自由变量，则方程（1）会有多个解）.

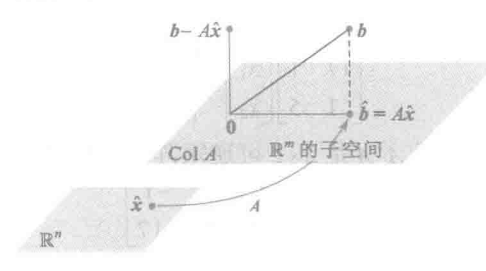


图 6-24 $\mathbb{R}^n$ 中的最小二乘解 $\hat{\boldsymbol{x}}$

若 $\hat{\boldsymbol{x}}$ 满足 $A\hat{\boldsymbol{x}}=\hat{\boldsymbol{b}}$，由 6.3 节的正交分解定理，投影 $\hat{\boldsymbol{b}}$ 具有性质 $(\boldsymbol{b}-\hat{\boldsymbol{b}})$ 与 Col A 正交，即 $(\boldsymbol{b}-A\hat{\boldsymbol{x}})$ 正交于 A 的每一列，如果 $\boldsymbol{a}_j$ 是 A 的任意列，那么 $\boldsymbol{a}_j\cdot(\boldsymbol{b}-A\hat{\boldsymbol{x}})=0$ 且 $\boldsymbol{a}_j{}^{\mathrm{T}}\cdot(\boldsymbol{b}-A\hat{\boldsymbol{x}})=0$，由于每一个 $\boldsymbol{a}_j^{\mathrm{T}}$ 是 A^{T} 的行，

$$A^{\mathrm{T}}(\boldsymbol{b}-A\hat{\boldsymbol{x}})=\boldsymbol{0} \tag{2}$$

（方程（2）也可由 6.1 节中的定理 3 推出.）故有

$$A^{\mathrm{T}}\boldsymbol{b}-A^{\mathrm{T}}A\hat{\boldsymbol{x}}=\boldsymbol{0}$$

$$A^{\mathrm{T}}A\hat{\boldsymbol{x}}=A^{\mathrm{T}}\boldsymbol{b}$$

此计算表明 $A\boldsymbol{x}=\boldsymbol{b}$ 的每个最小二乘解满足方程

$$A^{\mathrm{T}}A\boldsymbol{x}=A^{\mathrm{T}}\boldsymbol{b} \tag{3}$$

矩阵方程（3）表示的线性方程组常称为 $A\boldsymbol{x}=\boldsymbol{b}$ 的**法方程**，（3）的解通常用 $\hat{\boldsymbol{x}}$ 表示.

定理 13 *方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解集和法方程 $A^{\mathrm{T}}A\boldsymbol{x}=A^{\mathrm{T}}\boldsymbol{b}$ 的非空解集一致.*

证 我们已经说明最小二乘解集合是非空的，且这个 $\hat{\boldsymbol{x}}$ 满足法方程. 相反，假若 $\hat{\boldsymbol{x}}$ 满足 $A^{\mathrm{T}}A\hat{\boldsymbol{x}}=A^{\mathrm{T}}\boldsymbol{b}$，那么 $\hat{\boldsymbol{x}}$ 满足上面方程（2），从而说明 $\boldsymbol{b}-A\hat{\boldsymbol{x}}$ 与 A^{T} 的行正交，即与 A 的列向量正交. 因为 A 的列生成 $\mathrm{Col}\,A$，向量 $\boldsymbol{b}-A\hat{\boldsymbol{x}}$ 与所有 $\mathrm{Col}\,A$ 的向量正交，因此方程 $\boldsymbol{b}=A\hat{\boldsymbol{x}}+(\boldsymbol{b}-A\hat{\boldsymbol{x}})$ 是将 $\boldsymbol{b}$ 分解成 $\mathrm{Col}\,A$ 中的一个向量与正交于 $\mathrm{Col}\,A$ 的一个向量之和. 根据正交分解的唯一性，$A\hat{\boldsymbol{x}}$ 必须是将 $\boldsymbol{b}$ 投影到 $\mathrm{Col}\,A$ 的正交投影. 所以，$A\hat{\boldsymbol{x}}=\boldsymbol{b}$ 成立，且 $\hat{\boldsymbol{x}}$ 是一个最小二乘解. ■

例 1 求不相容方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解，其中

$$A=\begin{bmatrix}4&0\\0&2\\1&1\end{bmatrix},\qquad \boldsymbol{b}=\begin{bmatrix}2\\0\\11\end{bmatrix}$$

解 利用（3）计算

$$A^{\mathrm{T}}A=\begin{bmatrix}4&0&1\\0&2&1\end{bmatrix}\begin{bmatrix}4&0\\0&2\\1&1\end{bmatrix}=\begin{bmatrix}17&1\\1&5\end{bmatrix}$$

$$A^{\mathrm{T}}\boldsymbol{b}=\begin{bmatrix}4&0&1\\0&2&1\end{bmatrix}\begin{bmatrix}2\\0\\11\end{bmatrix}=\begin{bmatrix}19\\11\end{bmatrix}$$

那么方程 $A^{\mathrm{T}}A\boldsymbol{x}=A^{\mathrm{T}}\boldsymbol{b}$ 变成

$$\begin{bmatrix}17&1\\1&5\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}19\\11\end{bmatrix}$$

行变换可用于解此方程组，但由于 $A^{\mathrm{T}}A$ 是 2×2 可逆矩阵，很快计算得到

$$(A^{\mathrm{T}}A)^{-1}=\frac{1}{84}\begin{bmatrix}5&-1\\-1&17\end{bmatrix}$$

那么可解 $A^{\mathrm{T}}A\boldsymbol{x}=A^{\mathrm{T}}\boldsymbol{b}$ 如下：

$$\begin{aligned}\hat{\boldsymbol{x}}&=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b}\\&=\frac{1}{84}\begin{bmatrix}5&-1\\-1&17\end{bmatrix}\begin{bmatrix}19\\11\end{bmatrix}=\frac{1}{84}\begin{bmatrix}84\\168\end{bmatrix}=\begin{bmatrix}1\\2\end{bmatrix}\end{aligned}$$ ■

在许多计算中，$A^{\mathrm{T}}A$ 是可逆的，但并不总是这样，下面例子中的矩阵出现于统计学中的方差分析问题.

例 2 求 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解，其中

$$A=\begin{bmatrix}1&1&0&0\\1&1&0&0\\1&0&1&0\\1&0&1&0\\1&0&0&1\\1&0&0&1\end{bmatrix},\quad \boldsymbol{b}=\begin{bmatrix}-3\\-1\\0\\2\\5\\1\end{bmatrix}$$

$$A^{\mathrm{T}}A=\begin{bmatrix}1&1&1&1&1&1\\1&1&0&0&0&0\\0&0&1&1&0&0\\0&0&0&0&1&1\end{bmatrix}\begin{bmatrix}1&1&0&0\\1&1&0&0\\1&0&1&0\\1&0&1&0\\1&0&0&1\\1&0&0&1\end{bmatrix}=\begin{bmatrix}6&2&2&2\\2&2&0&0\\2&0&2&0\\2&0&0&2\end{bmatrix}$$

$$A^{\mathrm{T}}\boldsymbol{b}=\begin{bmatrix}1&1&1&1&1&1\\1&1&0&0&0&0\\0&0&1&1&0&0\\0&0&0&0&1&1\end{bmatrix}\begin{bmatrix}-3\\-1\\0\\2\\5\\1\end{bmatrix}=\begin{bmatrix}4\\-4\\2\\6\end{bmatrix}$$

矩阵方程 $A^{\mathrm{T}}A\boldsymbol{x}=A^{\mathrm{T}}\boldsymbol{b}$ 的增广矩阵是

$$\begin{bmatrix}6&2&2&2&4\\2&2&0&0&-4\\2&0&2&0&2\\2&0&0&2&6\end{bmatrix}\sim\begin{bmatrix}1&0&0&1&3\\0&1&0&-1&-5\\0&0&1&-1&-2\\0&0&0&0&0\end{bmatrix}$$

通解是：$x_1=3-x_4$，$x_2=-5+x_4$，$x_3=-2+x_4$，x_4 是自由变量. 所以，$A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘通解具有下面形式：

$$\hat{\boldsymbol{x}}=\begin{bmatrix}3\\-5\\-2\\0\end{bmatrix}+x_4\begin{bmatrix}-1\\1\\1\\1\end{bmatrix}$$ ■

下面的定理给出判断准则，在什么条件下，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解是唯一的.（当然，正交投影 $\hat{\boldsymbol{b}}$ 总是唯一的.）

定理 14 设 A 是 $m\times n$ 矩阵. 下面的条件是逻辑等价的：a）对于 $\mathbb{R}^m$ 中的每个 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 有唯一最小二乘解. b）A 的列是线性无关的. c）矩阵 $A^{\mathrm{T}}A$ 是可逆的.

当这些条件成立时，最小二乘解 $\hat{\boldsymbol{x}}$ 有下面的表示：

$$\hat{\boldsymbol{x}}=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b} \tag{4}$$

定理 14 证明的主要部分在习题 19~21 中给出，证明的同时也复习了第 4 章的概念. 用公式（4）计算 $\hat{\boldsymbol{x}}$ 的方法主要具有理论意义，当 $A^{\mathrm{T}}A$ 是 2×2 可逆矩阵时也可用手工计算.

当最小二乘解 $\hat{x}$ 用于产生 b 的近似逼近 $A\hat{x}$ 时，从 b 到 $A\hat{x}$ 的距离称为这个近似的**最小二乘误差**.

例 3　如例 1 给出 A 和 b，确定 $Ax=b$ 最小二乘解的最小二乘误差.

解　从例 1 可知，

$$b=\begin{bmatrix}2\\0\\11\end{bmatrix} \text{和}\ A\hat{x}=\begin{bmatrix}4&0\\0&2\\1&1\end{bmatrix}\begin{bmatrix}1\\2\end{bmatrix}=\begin{bmatrix}4\\4\\3\end{bmatrix}$$

因此

$$b-A\hat{x}=\begin{bmatrix}2\\0\\11\end{bmatrix}-\begin{bmatrix}4\\4\\3\end{bmatrix}=\begin{bmatrix}-2\\-4\\8\end{bmatrix}$$

$$\|b-A\hat{x}\|=\sqrt{(-2)^2+(-4)^2+8^2}=\sqrt{84}$$

最小二乘误差是 $\sqrt{84}$，对任意属于 $\mathbb{R}^2$ 的 x，从 b 到向量 Ax 的最小距离是 $\sqrt{84}$，见图 6-25，注意最小二乘解 $\hat{x}$ 自己并没有在图中出现.

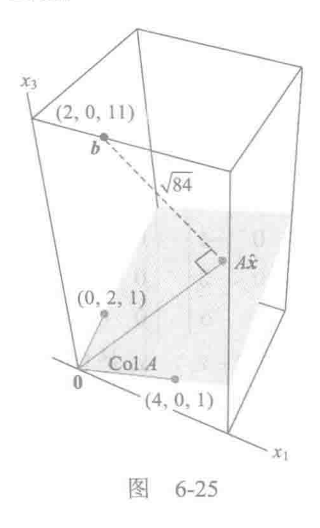


图　6-25

■

最小二乘解的另一个计算

下面的例子表明，当 A 的列向量不正交时，如何找到 $Ax=b$ 的最小二乘解，这类矩阵常出现在下节要讨论的线性回归问题中.

例 4　找出 $Ax=b$ 的最小二乘解，其中

$$A=\begin{bmatrix}1&-6\\1&-2\\1&1\\1&7\end{bmatrix},\quad b=\begin{bmatrix}-1\\2\\1\\6\end{bmatrix}$$

解　由于 A 的列 a_1 和列 a_2 相互正交，b 在 Col A 的正交投影如下：

$$\hat{\boldsymbol{b}} = \frac{\boldsymbol{b}\cdot\boldsymbol{a}_1}{\boldsymbol{a}_1\cdot\boldsymbol{a}_1}\cdot\boldsymbol{a}_1 + \frac{\boldsymbol{b}\cdot\boldsymbol{a}_2}{\boldsymbol{a}_2\cdot\boldsymbol{a}_2}\cdot\boldsymbol{a}_2 = \frac{8}{4}\boldsymbol{a}_1 + \frac{45}{90}\boldsymbol{a}_2 = \begin{bmatrix}2\\2\\2\\2\end{bmatrix} + \begin{bmatrix}-3\\-1\\1/2\\7/2\end{bmatrix} = \begin{bmatrix}-1\\1\\5/2\\11/2\end{bmatrix} \tag{5}$$

既然 $\hat{\boldsymbol{b}}$ 已知，我们可以解 $A\hat{\boldsymbol{x}} = \hat{\boldsymbol{b}}$. 这个很容易，因为我们已经知道 $\hat{\boldsymbol{b}}$ 用 A 的列线性表示时的权值.从（5）立刻得到：

$$\hat{\boldsymbol{x}} = \begin{bmatrix}8/4\\45/90\end{bmatrix} = \begin{bmatrix}2\\1/2\end{bmatrix}$$ ■

某些时候，最小二乘解问题的法方程可能是病态的，也就是 $A^{\mathrm{T}}A$ 中元素在计算中出现较小的误差，可导致解 $\hat{\boldsymbol{x}}$ 出现较大的误差. 如果 A 的列线性无关，最小二乘解常常可通过 A 的 QR 分解更可靠地求出（见 6.4 节的描述）.[⊖]

定理 15 给定一个 $m\times n$ 矩阵 A，且具有线性无关的列，取 $A=QR$ 是 A 类似定理 12 的 QR 分解，那么对每一个属于 $\mathbb{R}^m$ 的 $\boldsymbol{b}$，矩阵 $A\boldsymbol{x}=\boldsymbol{b}$ 有唯一的最小二乘解，其解为

$$\hat{\boldsymbol{x}} = R^{-1}Q^{\mathrm{T}}\boldsymbol{b} \tag{6}$$

证 取 $\hat{\boldsymbol{x}} = R^{-1}Q^{\mathrm{T}}\boldsymbol{b}$，那么 $A\hat{\boldsymbol{x}} = QR\hat{\boldsymbol{x}} = QRR^{-1}Q^{\mathrm{T}}\boldsymbol{b} = QQ^{\mathrm{T}}\boldsymbol{b}$.

由定理 12，Q 的列形成 $\mathrm{Col}\,A$ 的正交基，因此，由定理 10，$QQ^{\mathrm{T}}\boldsymbol{b}$ 是 $\boldsymbol{b}$ 在 $\mathrm{Col}\,A$ 上的正交投影 $\hat{\boldsymbol{b}}$，那么 $A\hat{\boldsymbol{x}} = \hat{\boldsymbol{b}}$ 说明 $\hat{\boldsymbol{x}}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解. $\hat{\boldsymbol{x}}$ 的唯一性可从定理 14 得出. ■

数值计算的注解 由于定理 15 中的 R 是上三角形矩阵，$\hat{\boldsymbol{x}}$ 可从方程

$$R\boldsymbol{x} = Q^{\mathrm{T}}\boldsymbol{b} \tag{7}$$

计算得到. 求解方程（7）时，通过回代过程或行变换比利用（6）计算 R^{-1} 更快.

例 5 求出 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解，其中

$$A = \begin{bmatrix}1&3&5\\1&1&0\\1&1&2\\1&3&3\end{bmatrix},\quad \boldsymbol{b} = \begin{bmatrix}3\\5\\7\\-3\end{bmatrix}$$

解 利用 6.4 节可得 A 的 QR 分解.

$$A = QR = \begin{bmatrix}1/2&1/2&1/2\\1/2&-1/2&-1/2\\1/2&-1/2&1/2\\1/2&1/2&-1/2\end{bmatrix}\begin{bmatrix}2&4&5\\0&2&3\\0&0&2\end{bmatrix}$$

[⊖] 在 G. Golub and C. Van Loan, *Matrix Computations*,3rd ed.(Baltimore: Johns Hopkins Press,1996), pp. 230-231 中给出 QR 分解方法和标准法方程方法的比较.

那么

$$Q^{\mathrm{T}}\boldsymbol{b}=\begin{bmatrix}1/2 & 1/2 & 1/2 & 1/2\\ 1/2 & -1/2 & -1/2 & 1/2\\ 1/2 & -1/2 & 1/2 & -1/2\end{bmatrix}\begin{bmatrix}3\\5\\7\\-3\end{bmatrix}=\begin{bmatrix}6\\-6\\4\end{bmatrix}$$

满足 $R\boldsymbol{x}=Q^{\mathrm{T}}\boldsymbol{b}$ 的最小二乘解是 $\hat{\boldsymbol{x}}$，也就是说，

$$\begin{bmatrix}2 & 4 & 5\\ 0 & 2 & 3\\ 0 & 0 & 2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}6\\-6\\4\end{bmatrix}$$

这个方程很容易解出，且 $\hat{\boldsymbol{x}}=\begin{bmatrix}10\\-6\\2\end{bmatrix}$. ■

练习题

1. 令 $A=\begin{bmatrix}1 & -3 & -3\\ 1 & 5 & 1\\ 1 & 7 & 2\end{bmatrix}$ 和 $\boldsymbol{b}=\begin{bmatrix}5\\-3\\-5\end{bmatrix}$，求 $A\boldsymbol{x}=\boldsymbol{b}$ 的一个最小二乘解，并且计算最小二乘解的误差.
2. 当 $\boldsymbol{b}$ 与 A 的列正交时，对 $A\boldsymbol{x}=\boldsymbol{b}$ 最小二乘问题的解，你可得出什么结论？

习题 6.5

在习题 1~4 中，求 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘问题的解.（a）通过构造法方程求 $\hat{\boldsymbol{x}}$，（b）直接解 $\hat{\boldsymbol{x}}$.

1. $A=\begin{bmatrix}-1 & 2\\ 2 & -3\\ -1 & 3\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}4\\1\\2\end{bmatrix}$
2. $A=\begin{bmatrix}2 & 1\\ -2 & 0\\ 2 & 3\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}-5\\8\\1\end{bmatrix}$
3. $A=\begin{bmatrix}1 & -2\\ -1 & 2\\ 0 & 3\\ 2 & 5\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}3\\1\\-4\\2\end{bmatrix}$
4. $A=\begin{bmatrix}1 & 3\\ 1 & -1\\ 1 & 1\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}5\\1\\0\end{bmatrix}$

在习题 5~6 中，求方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的所有最小二乘解.

5. $A=\begin{bmatrix}1 & 1 & 0\\ 1 & 1 & 0\\ 1 & 0 & 1\\ 1 & 0 & 1\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}1\\3\\8\\2\end{bmatrix}$
6. $A=\begin{bmatrix}1 & 1 & 0\\ 1 & 1 & 0\\ 1 & 1 & 0\\ 1 & 0 & 1\\ 1 & 0 & 1\\ 1 & 0 & 1\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}7\\2\\3\\6\\5\\4\end{bmatrix}$
7. 计算习题 3 中最小二乘问题相关的最小二乘误差.
8. 计算习题 4 中最小二乘问题相关的最小二乘误差.

在习题 9~12 中，求（a）$\boldsymbol{b}$ 在 Col A 的正交投影，（b）最小二乘问题 $A\boldsymbol{x}=\boldsymbol{b}$ 的解.

9. $A=\begin{bmatrix}1&5\\3&1\\-2&4\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}4\\-2\\-3\end{bmatrix}$

10. $A=\begin{bmatrix}1&2\\-1&4\\1&2\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}3\\-1\\5\end{bmatrix}$

11. $A=\begin{bmatrix}4&0&1\\1&-5&1\\6&1&0\\1&-1&-5\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}9\\0\\0\\0\end{bmatrix}$

12. $A=\begin{bmatrix}1&1&0\\1&0&-1\\0&1&1\\-1&1&-1\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}2\\5\\6\\6\end{bmatrix}$

13. 令 $A=\begin{bmatrix}3&4\\-2&1\\3&4\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}11\\-9\\5\end{bmatrix}$，$\boldsymbol{u}=\begin{bmatrix}5\\-1\end{bmatrix}$ 和 $\boldsymbol{v}=\begin{bmatrix}5\\-2\end{bmatrix}$，计算 $A\boldsymbol{u}$ 和 $A\boldsymbol{v}$，并与 $\boldsymbol{b}$ 相比较，$\boldsymbol{u}$ 是否为 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解？（不用计算最小二乘问题的解，回答这个问题.）

14. 令 $A=\begin{bmatrix}2&1\\-3&-4\\3&2\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}5\\4\\4\end{bmatrix}$，$\boldsymbol{u}=\begin{bmatrix}4\\-5\end{bmatrix}$ 和 $\boldsymbol{v}=\begin{bmatrix}6\\-5\end{bmatrix}$，计算 $A\boldsymbol{u}$ 和 $A\boldsymbol{v}$，并与 $\boldsymbol{b}$ 相比较，是否有可能至少有一个 $\boldsymbol{u}$ 或 $\boldsymbol{v}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解？（不用计算最小二乘问题的解，回答这个问题.）

在习题 15~16 中，利用分解 $A=QR$ 求 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘问题的解.

15. $A=\begin{bmatrix}2&3\\2&4\\1&1\end{bmatrix}=\begin{bmatrix}2/3&-1/3\\2/3&2/3\\1/3&-2/3\end{bmatrix}\begin{bmatrix}3&5\\0&1\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}7\\3\\1\end{bmatrix}$

16. $A=\begin{bmatrix}1&-1\\1&4\\1&-1\\1&4\end{bmatrix}=\begin{bmatrix}1/2&-1/2\\1/2&1/2\\1/2&-1/2\\1/2&1/2\end{bmatrix}\begin{bmatrix}2&3\\0&5\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}-1\\6\\5\\7\end{bmatrix}$

在习题 17~18 中，A 是 $m\times n$ 矩阵且 $\boldsymbol{b}$ 属于 $\mathbb{R}^m$，判断下列命题正确或错误，验证每一个答案.

17. a. 一般最小二乘问题是求出 $\boldsymbol{x}$ 使得 $A\boldsymbol{x}$ 尽可能接近 $\boldsymbol{b}$.
 b. 方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解是适合方程 $A\hat{\boldsymbol{x}}=\hat{\boldsymbol{b}}$ 的向量 $\hat{\boldsymbol{x}}$，此处 $\hat{\boldsymbol{b}}$ 是 $\boldsymbol{b}$ 在 Col A 上的正交投影.
 c. 方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解是向量 $\hat{\boldsymbol{x}}$，对所有属于 $\mathbb{R}^n$ 的 $\boldsymbol{x}$ 满足
$$\|\boldsymbol{b}-A\boldsymbol{x}\|\leqslant\|\boldsymbol{b}-A\hat{\boldsymbol{x}}\|$$
 d. 方程 $A^{\mathrm{T}}A\boldsymbol{x}=A^{\mathrm{T}}\boldsymbol{b}$ 的任意解是方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解.
 e. 如果 A 的列线性无关，那么方程 $A\boldsymbol{x}=\boldsymbol{b}$ 只有一个最小二乘解.

18. a. 如果 $\boldsymbol{b}$ 属于 A 的列空间，那么方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的每个解都是最小二乘解.
 b. 方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解，是属于 A 的列空间中最接近 $\boldsymbol{b}$ 的点.
 c. 方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解是一系列的分量，当它们作用在 A 的列，产生 $\boldsymbol{b}$ 在 Col A 上的正交投影.
 d. 如果 $\hat{\boldsymbol{x}}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的一个最小二乘解，那么 $\hat{\boldsymbol{x}}=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b}$.
 e. 法方程计算最小二乘解的方法总是可靠的.
 f. 如果 A 有一个 QR 分解，如 $A=QR$，那么求 $A\boldsymbol{x}=\boldsymbol{b}$ 最小二乘解的最好方法是计算 $\hat{\boldsymbol{x}}=R^{-1}Q^{\mathrm{T}}\boldsymbol{b}$.

19. 设 A 是一个 $m\times n$ 矩阵，利用下面的步骤说明，$\boldsymbol{x}$ 属于 $\mathbb{R}^n$ 且满足 $A\boldsymbol{x}=\boldsymbol{0}$ 的充分必要条件是 $A^{\mathrm{T}}A\boldsymbol{x}=\boldsymbol{0}$. 这将证明 Nul $A=$ Nul $A^{\mathrm{T}}A$.
 a. 证明：如果 $A\boldsymbol{x}=\boldsymbol{0}$，那么 $A^{\mathrm{T}}A\boldsymbol{x}=\boldsymbol{0}$.
 b. 假若 $A^{\mathrm{T}}A\boldsymbol{x}=\boldsymbol{0}$，解释为什么 $\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}A\boldsymbol{x}=0$ 且利用此式说明 $A\boldsymbol{x}=\boldsymbol{0}$.

20. 设 A 是一个 $m\times n$ 矩阵且 $A^{\mathrm{T}}A$ 是可逆的，证明

A 的列向量线性无关.(注意：不能假设 A 可逆，且更不一定是方阵.)

21. 设 A 是 $m\times n$ 矩阵且列向量线性无关(注意：A 不一定是方阵).
    a. 利用习题 19 证明 A^TA 是可逆矩阵.
    b. 解释为什么 A 具有至少与行一样多的列.
    c. 确定 A 的秩.
22. 利用习题 19 证明 $\text{rank}\, A^TA=\text{rank}\, A$.（提示：$A^TA$ 有多少列？如何与 A^TA 的秩联系起来？）
23. 假设 A 是 $m\times n$ 矩阵且有线性无关的列，$\boldsymbol{b}$ 属于 $\mathbb{R}^m$，利用法方程给一个计算 $\hat{\boldsymbol{b}}$ 的公式，$\hat{\boldsymbol{b}}$ 是 $\boldsymbol{b}$ 在 $\text{Col}\, A$ 上的投影.(提示：首先找到 $\hat{\boldsymbol{x}}$，这个公式不需要 $\text{Col}\, A$ 的正交基.)
24. 当 A 的列是单位正交时，找出方程 $A\boldsymbol{x}=\boldsymbol{b}$ 最小二乘解的一个公式.
25. 列出下列方程组的所有最小二乘解.

$$x+y=2$$
$$x+y=4$$

26. [M] 4.8 节的例 3 显示一个低通线性滤波器，将信号 $\{y_k\}$ 过滤为 $\{\boldsymbol{y}_{k+1}\}$. 并将高频信号 $\{w_k\}$ 过滤为零信号，其中 $y_k=\cos(\pi k/4)$ 和 $w_k=\cos(3\pi k/4)$，下面计算将设计一个滤波器近似满足此特性. 滤波器方程为

$$a_0y_{k+2}+a_1y_{k+1}+a_2y_k=z_k\text{，对所有 } k \text{ 成立} \quad (8)$$

因为信号是周期的，且周期为 8，对方程(8)只需研究 $k=0,\cdots,7$ 即可. 滤波器的作用是将上面描述的两个信号转化为两组 8 个方程组.

$$\begin{array}{c} \\ k=0 \\ k=1 \\ \vdots \\ \\ \\ \\ \\ k=7 \end{array}
\begin{array}{c} \begin{array}{ccc} y_{k+2} & y_{k+1} & y_k \end{array} \\ \begin{bmatrix} 0 & 0.7 & 1 \\ -0.7 & 0 & 0.7 \\ -1 & -0.7 & 0 \\ -0.7 & -1 & -0.7 \\ 0 & -0.7 & -1 \\ 0.7 & 0 & -0.7 \\ 1 & 0.7 & 0 \\ 0.7 & 1 & 0.7 \end{bmatrix} \end{array}
\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} =
\begin{array}{c} y_{k+1} \\ \begin{bmatrix} 0.7 \\ 0 \\ -0.7 \\ -1 \\ -0.7 \\ 0 \\ 0.7 \\ 1 \end{bmatrix} \end{array}$$

$$\begin{array}{c} \\ k=0 \\ k=1 \\ \vdots \\ \\ \\ \\ \\ k=7 \end{array}
\begin{array}{c} \begin{array}{ccc} w_{k+2} & w_{k+1} & w_k \end{array} \\ \begin{bmatrix} 0 & -0.7 & 1 \\ 0.7 & 0 & -0.7 \\ -1 & 0.7 & 0 \\ 0.7 & -1 & 0.7 \\ 0 & 0.7 & -1 \\ -0.7 & 0 & 0.7 \\ 1 & -0.7 & 0 \\ -0.7 & 1 & -0.7 \end{bmatrix} \end{array}
\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} =
\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

写出一个方程 $A\boldsymbol{x}=\boldsymbol{b}$，其中 A 是一个 16×3 矩阵，由上面两个系数矩阵组成，$\boldsymbol{b}$ 属于 $\mathbb{R}^{16}$，由上面两个方程右边的向量组成. 计算 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解确定的 a_0,a_1,a_2（上面数据中的 0.7 是 $\sqrt{2}/2$ 的一个近似值，用以说明典型的实际问题中如何进行计算.如果用 0.707 代替，所得过滤系数与精确的算术计算结果 $\sqrt{2}/4$，$1/2$，$\sqrt{2}/4$ 相比至少有 7 位相同的十进制位数）.

练习题答案

1. 首先计算

$$A^TA=\begin{bmatrix} 1 & 1 & 1 \\ -3 & 5 & 7 \\ -3 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{bmatrix}=\begin{bmatrix} 3 & 9 & 0 \\ 9 & 83 & 28 \\ 0 & 28 & 14 \end{bmatrix}$$

$$A^T\boldsymbol{b}=\begin{bmatrix} 1 & 1 & 1 \\ -3 & 5 & 7 \\ -3 & 1 & 2 \end{bmatrix}\begin{bmatrix} 5 \\ -3 \\ -5 \end{bmatrix}=\begin{bmatrix} -3 \\ -65 \\ -28 \end{bmatrix}$$

下一步，行化简标准方程 $A^{\mathrm{T}}A\boldsymbol{x}=A^{\mathrm{T}}\boldsymbol{b}$ 对应的增广矩阵：

$$\begin{bmatrix}3&9&0&-3\\9&83&28&-65\\0&28&14&-28\end{bmatrix}\sim\begin{bmatrix}1&3&0&-1\\0&56&28&-56\\0&28&14&-28\end{bmatrix}\sim\cdots\sim\begin{bmatrix}1&0&-3/2&2\\0&1&1/2&-1\\0&0&0&0\end{bmatrix}$$

一般最小二乘解是 $x_1=2+\dfrac{3}{2}x_3$， $x_2=-1-\dfrac{1}{2}x_3$， x_3 是自由变量. 取 $x_3=0$，可得一个特殊解：

$$\hat{\boldsymbol{x}}=\begin{bmatrix}2\\-1\\0\end{bmatrix}$$

为得到最小二乘误差，计算

$$\hat{\boldsymbol{b}}=A\hat{\boldsymbol{x}}=\begin{bmatrix}1&-3&-3\\1&5&1\\1&7&2\end{bmatrix}\begin{bmatrix}2\\-1\\0\end{bmatrix}=\begin{bmatrix}5\\-3\\-5\end{bmatrix}$$

结果出现 $\hat{\boldsymbol{b}}=\boldsymbol{b}$，所以 $\|\boldsymbol{b}-\hat{\boldsymbol{b}}\|=0$，由于 $\boldsymbol{b}$ 恰好属于 Col A，最小二乘误差为零.

2. 如果 $\boldsymbol{b}$ 与 A 的列正交，那么 $\boldsymbol{b}$ 在 A 的列空间 A 的投影是 $\boldsymbol{0}$，在这种情形下， $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解 $\hat{\boldsymbol{x}}$ 满足 $A\hat{\boldsymbol{x}}=\boldsymbol{0}$.

6.6 线性模型中的应用

科学和工程中的一项任务就是分析或理解一些数量变化之间的联系，本节描述各种情形下，数据被用作构造或验证一个公式，该公式可预测一个变量作为其他变量值的函数. 在每种情形下，问题会等同于求解一个最小二乘问题.

为了更容易应用所讨论的实际问题，如读者以后学习遇到的专业知识，我们选取科学和工程数据最常见的统计分析记号. 将 $A\boldsymbol{x}=\boldsymbol{b}$ 写成 $X\boldsymbol{\beta}=\boldsymbol{y}$，且称 X 为**设计矩阵**， $\boldsymbol{\beta}$ 为**参数向量**， $\boldsymbol{y}$ 为**观测向量**.

最小二乘直线

变量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 之间最简单的关系是线性方程： $\boldsymbol{y}=\beta_0+\beta_1\boldsymbol{x}$.[㊀] 实验中数据常常给出点列 $(x_1,y_1),\cdots,(x_n,y_n)$，而它们的图形近似接近于直线，我们希望确定参数 β_0 和 β_1，使得直线尽可能“接近”这些点.

假若 β_0 和 β_1 固定，考虑图 6-26 的直线 $\boldsymbol{y}=\beta_0+\beta_1\boldsymbol{x}$. 对应每一个数据点 (x_j,y_j)，有一个在直线上点列 $(x_j,\beta_0+\beta_1x_j)$ 具有同样的 x 坐标. 我们称 $\boldsymbol{y}_j$ 为 $\boldsymbol{y}$ 的观测值，而 $\beta_0+\beta_1x_j$ 为 $\boldsymbol{y}$ 的预测值（由直线确定），观测 $\boldsymbol{y}$ 值和预测 $\boldsymbol{y}$ 值的差称为余差.

㊀ 这个记号在最小二乘直线中代替 $\boldsymbol{y}=m\boldsymbol{x}+\boldsymbol{b}$.

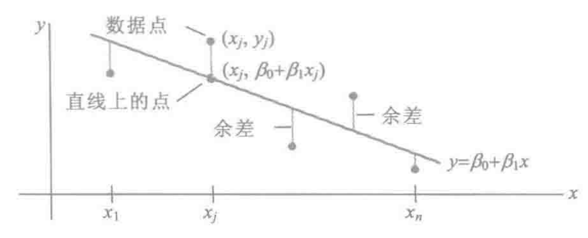


图 6-26 实验数据的直线拟合

有几种方法来度量直线如何“接近”数据，最常见的选择是余差平方之和（主要原因是数学计算简单），**最小二乘直线** $y=\beta_0+\beta_1 x$ 是余差平方之和最小，这条直线也称为 $\boldsymbol{y}$ 对 $\boldsymbol{x}$ **的回归直线**，因为假设数据中的任何误差只出现在 y 坐标，直线的系数 β_0,β_1 被称为（线性）**回归系数**.[⊖]

如果数据点在直线上，参数 β_0 和 β_1 满足方程

$$\begin{array}{ccc} \text{预测的} & & \text{观测的} \\ \boldsymbol{y}\text{ 值} & & \boldsymbol{y}\text{ 值} \\ \beta_0+\beta_1 x_1 & = & y_1 \\ \beta_0+\beta_1 x_2 & = & y_2 \\ \vdots & & \vdots \\ \beta_0+\beta_1 x_n & = & y_n \end{array}$$

我们可将这个方程写成

$$X\boldsymbol{\beta}=\boldsymbol{y}\text{，其中 } X=\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix},\quad \boldsymbol{\beta}=\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix},\quad \boldsymbol{y}=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{1}$$

当然，如果数据点不在直线上，就没有参数 β_0,β_1 使得 $X\boldsymbol{\beta}$ 中的预测 $\boldsymbol{y}$ 值与 $\boldsymbol{y}$ 中的观测 $\boldsymbol{y}$ 值相等，而 $X\boldsymbol{\beta}=\boldsymbol{y}$ 没有解. 这就是 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解问题，只是记法不同！

向量 $X\boldsymbol{\beta}$ 与 $\boldsymbol{y}$ 距离的平方精确表达为余差的平方之和，于是，使平方和最小的 $\boldsymbol{\beta}$ 同样使得 $X\boldsymbol{\beta}$ 与 $\boldsymbol{y}$ 之间的距离最小，计算 $X\boldsymbol{\beta}=\boldsymbol{y}$ 的最小二乘问题等价于找出 $\boldsymbol{\beta}$，它确定图 6-26 中的最小二乘直线.

例 1 求方程 $y=\beta_0+\beta_1 x$ 的最小二乘直线，最佳拟合数据点为 (2,1)，(5,2)，(7,3)，(8,3).

解 利用数据的 x 坐标，构造 (1) 中的矩阵 X 和 y 坐标构造向量 $\boldsymbol{y}$：

$$X=\begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix},\quad \boldsymbol{y}=\begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}$$

对 $X\boldsymbol{\beta}=\boldsymbol{y}$ 的最小二乘解，得到法方程（用新记号）：

⊖ 如果测量的余差是 x，而不是 y. 在描点和计算回归直线之前只需交换数据 (x_j, y_j) 的坐标，如果两个坐标都有误差，你必须选择直线，使得数据点到直线的正交（垂直）距离平方和最小，见 7.5 节的练习题.

$$X^{\mathrm{T}}X\boldsymbol{\beta}=X^{\mathrm{T}}\boldsymbol{y}$$

也就是说，计算

$$X^{\mathrm{T}}X=\begin{bmatrix}1&1&1&1\\2&5&7&8\end{bmatrix}\begin{bmatrix}1&2\\1&5\\1&7\\1&8\end{bmatrix}=\begin{bmatrix}4&22\\22&142\end{bmatrix}$$

$$X^{\mathrm{T}}\boldsymbol{y}=\begin{bmatrix}1&1&1&1\\2&5&7&8\end{bmatrix}\begin{bmatrix}1\\2\\3\\3\end{bmatrix}=\begin{bmatrix}9\\57\end{bmatrix}$$

标准方程是

$$\begin{bmatrix}4&22\\22&142\end{bmatrix}\begin{bmatrix}\beta_0\\\beta_1\end{bmatrix}=\begin{bmatrix}9\\57\end{bmatrix}$$

因此

$$\begin{bmatrix}\beta_0\\\beta_1\end{bmatrix}=\begin{bmatrix}4&22\\22&142\end{bmatrix}^{-1}\begin{bmatrix}9\\57\end{bmatrix}=\frac{1}{84}\begin{bmatrix}142&-22\\-22&4\end{bmatrix}\begin{bmatrix}9\\57\end{bmatrix}=\frac{1}{84}\begin{bmatrix}24\\30\end{bmatrix}=\begin{bmatrix}2/7\\5/14\end{bmatrix}$$

这样，最小二乘直线的方程为 $y=\frac{2}{7}+\frac{5}{14}x$，见图 6-27.

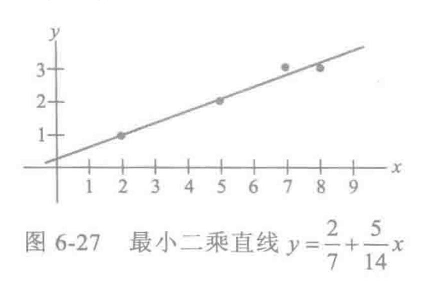


图 6-27　最小二乘直线 $y=\frac{2}{7}+\frac{5}{14}x$ ■

在计算最小二乘直线之前，常见的练习是计算原来 x 值的平均 $\bar{x}$，并形成一个新坐标 $x^{*}=x-\bar{x}$. 新的 x 坐标被称为**平均偏差形式**. 在这种情形下，设计矩阵的两列是正交的，像 6.5 节的例 4，标准方程的解是简化的，见习题 17 和习题 18.

一般线性模型

在一些应用中，必须将数据点拟合为非直线形式，在下面的例子中，矩阵方程仍然是 $X\boldsymbol{\beta}=\boldsymbol{y}$，但不同问题的 X 会从一个问题变到下一个，统计学家常引入余差向量 $\boldsymbol{\varepsilon}$，定义为 $\boldsymbol{\varepsilon}=\boldsymbol{y}-X\boldsymbol{\beta}$，并且记作

$$\boldsymbol{y}=X\boldsymbol{\beta}+\boldsymbol{\varepsilon}$$

任何具有这种形式的方程称为**线性模型**. 一旦 X 和 $\boldsymbol{y}$ 被确定，为使 $\boldsymbol{\varepsilon}$ 长度达到最小化，相当于找出 $X\boldsymbol{\beta}=\boldsymbol{y}$ 的最小二乘解，在每种情形下，最小二乘问题解 $\hat{\boldsymbol{\beta}}$ 是下面标准方程的解.

$$X^{\mathrm{T}}X\boldsymbol{\beta}=X^{\mathrm{T}}\boldsymbol{y}$$

其他曲线的最小二乘拟合

当“分散画出”的数据点 $(x_1, y_1), \cdots, (x_n, y_n)$ 不接近任何直线时，一个合适的假定是 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 具有其他函数关系.

下面三个例子说明，如何将数据拟合为如下一般形式的曲线.

$$y=\beta_0 f_0(x)+\beta_1 f_1(x)+\cdots+\beta_k f_k(x) \tag{2}$$

此处 $f_0, \cdots, f_k$ 是已知函数，$\beta_0, \cdots, \beta_k$ 是待定参数. 下面将看到，方程（2）描述一个线性模型，因为它是未知参数的线性模型.

对特殊的 x 值，（2）式给出 y 的预测或“拟合”值，观测值与预测值之间的差为余差，参数 $\beta_0, \cdots, \beta_k$ 的确定需满足余差平方之和最小.

例 2 若数据点 $(x_1, y_1), \cdots, (x_n, y_n)$ 明显位于某条抛物线之上，而不是一条直线上. 例如，如果 x 轴表示某公司的产量水平，而 y 轴表示生产水平为每天 x 单位时的平均费用，那么一个典型的成本曲线看起来像开口向上的抛物线（见图 6-28），在生态系统中，一个开口向下的抛物线常用于对一种植物中营养成分的净初始产量的建模，它是树叶表面面积的函数（见图 6-29）. 若我们用下列形式的方程逼近数据

$$y=\beta_0+\beta_1 x+\beta_2 x^2 \tag{3}$$

方程（3）给出描述数据的“最小二乘拟合”对应的线性模型.

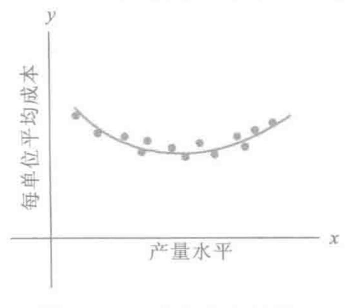


图 6-28 平均成本曲线

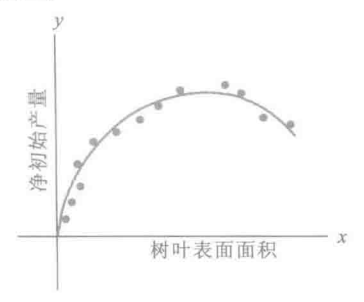


图 6-29 营养成分的产量

解 理想的关系用方程（3）描述，假若实际的参数值为 β_0, β_1, β_2，那么第一个数据点 (x_1, y_1) 适合下列形式的方程：

$$y_1=\beta_0+\beta_1 x_1+\beta_2 x_1^2+\varepsilon_1$$

此处 ε_1 是观测值 y_1 和预测值 y 值 $\beta_0+\beta_1 x_1+\beta_2 x_1^2$ 的余差，对每一个点集，我们可写出类似的方程：

$$\begin{aligned}
y_1&=\beta_0+\beta_1 x_1+\beta_2 x_1^2+\varepsilon_1\\
y_2&=\beta_0+\beta_1 x_2+\beta_2 x_2^2+\varepsilon_2\\
&\vdots\\
y_n&=\beta_0+\beta_1 x_n+\beta_2 x_n^2+\varepsilon_n
\end{aligned}$$

可将上述方程组简单描述为 $\boldsymbol{y}=X\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ 的形式，通过检查方程的前面几行和观察数据形状，我们可以求出 X.

$$\underset{\boldsymbol{y}}{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}} = \underset{X}{\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix}} \underset{\boldsymbol{\beta}}{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}} + \underset{\boldsymbol{\varepsilon}}{\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}}$$

■

例 3 如果数据点具有如图 6-30 的模式，那么一个合适的模型是下面形式的方程：

$$y=\beta_0+\beta_1 x+\beta_2 x^2+\beta_3 x^3$$

这类数据可能来自公司的总成本，描述的是一个关于产量水平的函数，求线性模型并用最小二乘拟合这类数据 $(x_1,y_1),\cdots,(x_n,y_n)$.

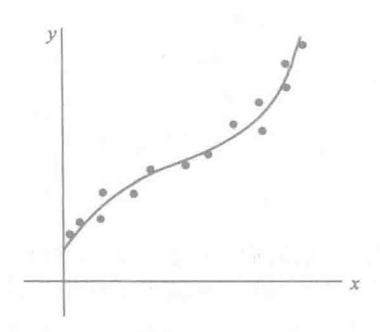


图 6-30 沿立方曲线的数据点

解 类似例 2 的分析，我们得到

$$\underset{\text{观测向量}}{\boldsymbol{y}=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}, \quad \underset{\text{设计矩阵}}{X=\begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & x_n^3 \end{bmatrix}}, \quad \underset{\text{参数向量}}{\boldsymbol{\beta}=\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}}, \quad \underset{\text{余差向量}}{\boldsymbol{\varepsilon}=\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}}$$

■

多重回归

假若一个实验包含两个独立变量（例如 u 和 v）和一个函数变量（例如 y）. 一个简单的通过 u 和 v 来预测 y 的方程有如下形式：

$$y=\beta_0+\beta_1 u+\beta_2 v \tag{4}$$

更一般的预测方程具有下面的形式：

$$y=\beta_0+\beta_1 u+\beta_2 v+\beta_3 u^2+\beta_4 uv+\beta_5 v^2 \tag{5}$$

这个方程常用于地质学，例如，模拟地面侵蚀、冰川、土壤酸碱性以及其他数据. 这种情形的最小二乘拟合称为*趋势曲面*.

方程（4）和（5）都可以推出一个线性模型，因为它们是未知参数的线性关系（尽管 u 和

v 是乘法）. 一般地，一个线性模型是指 y 可由下面方程来预测：

$$y = \beta_0 f_0(u,v) + \beta_1 f_1(u,v) + \cdots + \beta_k f_k(u,v)$$

此处，$f_0,\cdots,f_k$ 是某类已知函数，$\beta_0,\cdots,\beta_k$ 是未知权值.

例 4 在地理学中，局部地形模型由数据 $(u_1,v_1,y_1),\cdots,(u_n,v_n,y_n)$ 来构造，此处 u_j,v_j,y_j 分别是地形的纬度、经度和高度，描述基于方程（4）的线性模型且给出这些数据的最小二乘拟合，该解称为最小二乘平面. 见图 6-31.

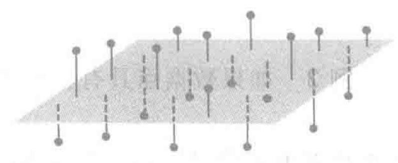


图 6-31 最小二乘平面

解 我们希望数据满足下列方程：

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 u_1 + \beta_2 v_1 + \varepsilon_1 \\ y_2 &= \beta_0 + \beta_1 u_2 + \beta_2 v_2 + \varepsilon_2 \\ &\vdots \qquad\qquad \vdots \\ y_n &= \beta_0 + \beta_1 u_n + \beta_2 v_n + \varepsilon_n \end{aligned}$$

这个方程组的矩阵形式是 $\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$，此处

$$\underset{\text{观测值}}{\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}},\quad \underset{\text{设计矩阵}}{X = \begin{bmatrix} 1 & u_1 & v_1 \\ 1 & u_2 & v_2 \\ \vdots & \vdots & \vdots \\ 1 & u_n & v_n \end{bmatrix}},\quad \underset{\text{参数向量}}{\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}},\quad \underset{\text{余差向量}}{\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}}$$

■

例 4 表明，多重回归的线性模型和前面例题中的简单回归模型具有同样的抽象形式，线性代数为我们理解所有线性模型内在的一般原理提供了帮助，定义只要 X 适当，关于 $\boldsymbol{\beta}$ 的标准方程就具有相同的矩阵形式，不管包含多少变量. 这样，对 $X^{\mathrm{T}}X$ 可逆的任何线性模型，最小二乘中的 $\hat{\boldsymbol{\beta}}$ 总可由 $(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}\boldsymbol{y}$ 计算得到.

进一步阅读

Ferguson, J., *Introduction to Linear Algebra in Geology* (New York: Chapman & Hall, 1994).

Krumbein, W. C., and F. A. Graybill, *An Introduction to Statistical Models in Geology* (New York: McGraw-Hill, 1965).

Legendre, P., and L. Legendre, *Numerical Ecology* (Amsterdam: Elsevier, 1998).

Unwin, David J., *An Introduction to Trend Surface Analysis*, Concepts and Techniques in Modern Geography, No. 5 (Norwich, England: Geo Books, 1975).

练习题

某产品的月销售额受季节波动影响，近似销售数据曲线具有以下形式：

$$y = \beta_0 + \beta_1 x + \beta_2 \sin(2\pi x/12)$$

这里 x 是按月统计的时间，$\beta_0 + \beta_1 x$ 给出基本销售趋势，其中的正弦项反映季节对销售的影响，给出上面最小二乘拟合方程中，线性模型的设计矩阵和参数向量，假设数据是 $(x_1,y_1),\cdots,(x_n,y_n)$.

习题 6.6

在习题 1~4 中，求出最小二乘直线方程 $y=\beta_0+\beta_1 x$，其为给定下数据点的最佳拟合.

1. （0,1），（1,1），（2,2），（3,2）
2. （1,0），（2,1），（4,2），（5,3）
3. （−1,0），（0,1），（1,2），（2,4）
4. （2,3），（3,2），（5,1），（6,0）
5. 假设 X 是用最小二乘直线拟合数据 (x_1,y_1)，$\cdots,(x_n,y_n)$ 对应得到的设计矩阵，利用 6.5 节的定理证明，标准方程具有唯一解的充分必要条件是：数据中至少有两个数据点具有不同的 x 坐标.
6. 若 X 是例 2 中最小二乘拟合抛物数据点 (x_1,y_1)，$\cdots,(x_n,y_n)$ 对应的一个设计矩阵，若 x_1,x_2,x_3 是不同的数据，解释在最小二乘意义下，为什么只有一个最佳抛物拟合数据（见习题 5）.
7. 一个实验产生的数据为(1, 1.8),(2,2.7),(3,3.4),（4,3.8）和（5,3.9），描述用下列函数形式生成的最小二乘拟合模型：
$$y=\beta_1 x+\beta_2 x^2$$
这种函数会出现在这样的情形下，即当销售的总量影响产品的价格设定时，销售 x 单位产品的收益计算中.
   a. 给出设计矩阵、观测向量和未知参数向量.
   b. [M]找出数据对应的最小二乘曲线.
8. 一个简单描述某公司花费变量的模型曲线，作为销售水平 x 时的函数，有如下形式：
$$y=\beta_1 x+\beta_2 x^2+\beta_3 x^3$$
由于不包含固定花费，所以没有常数项.
   a. 对于数据 $(x_1,y_1),\cdots,(x_n,y_n)$，给出线性模型的最小二乘拟合方程对应的设计矩阵和参数向量.
   b. [M]求出最小二乘曲线拟合数据：（4,1.58），（6,2.08），（8,2.5），（10,2.8），（12,3.1），（14,3.4），（16,3.8），（18,4.32），有数千个数据，如果可能，画图说明数据点和立方近似的图形.
9. 某一实验得到的数据为（1,7.9），（2,5.4）和（3,−0.9），描述由下列形式的函数拟合这些数据产生的最小二乘模型：
$$y=A\cos x+B\sin x$$
10. 若放射性物质 A 和 B 分别具有衰变常数 0.02 和 0.07，如果一个含两种物质的混合物，在时刻 $t=0$ 时，包含 A 物质 M_A 克和 B 物质 M_B 克，那么在时刻 t，混合物中总量 y 的模型是：
$$y=M_{\mathrm{A}}\mathrm{e}^{-0.02t}+M_{\mathrm{B}}\mathrm{e}^{-0.07t} \qquad (6)$$
若初始含量 M_{A} 和 M_{B} 未知，但在几个时刻，科学家可以测得并记录到下列时刻和含量的数据 (t_i,y_i)：（10, 21.34），（11, 20.68），（12, 20.05），（14, 18.87）和（15, 18.30）.
   a. 描述可以估计 M_{A} 和 M_{B} 的线性模型.
   b. [M]找出基于方程（6）的最小二乘曲线.
11. [M]根据开普勒第一定律，一个彗星（见图 6-32）应该有椭圆、抛物线和双曲轨道（当行星的引力可以忽略时）. 对合适的极坐标，一个彗星的位置 (r,θ) 适合下面形式的方程：
$$r=\beta+e(r.\cos\vartheta)$$
此处 β 是常数，e 是轨道的离心率，当 $0\leqslant e<1$ 时对应椭圆，$e=1$ 时对应抛物线，$e>1$ 时对应双曲线，若新观测出现的彗星有下列数据，确定轨道的类型并预测当 $\vartheta=4.6$（弧度）时彗星的位置.[⊖]

ϑ	0 .88	1.10	1.42	1.77	2.14
r	3.00	2.30	1.65	1.25	1.01

⊖ 用最小二乘解拟合数据的基本思想来自 k. F. Gauss（同时还有 A. Legendre 的独立工作）. 最早是在 1801 年，他用这种方法确定星状的谷神星的轨迹，当谷神星被发现 40 天后，它消失在太阳后面，Gauss 预测 10 个月之后它会重新出现且他给出谷神星的具体位置，精确的预测震惊了欧洲的科学界.

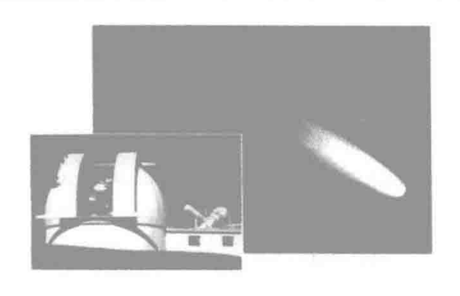


图 6-32　上一次出现于 1986 年的哈雷彗星，它将再次出现于 2061 年

12. [M]健康儿童的心脏收缩压 p（毫米水银柱）和体重 w（镑）之间的近似关系满足方程

$$\beta_0 + \beta_1 \ln w = p$$

利用下面实验数据估计健康儿童体重为 100 镑时的心脏收缩压.

w	44	61	81	113	131
$\ln w$	3.78	4.11	4.41	4.73	4.88
p	91	98	103	110	112

13. [M]为测量飞机起飞表演，飞机的水平位置从 $t=0$ 到 $t=12$ 每秒测量一次，具体位置（英尺）是：0, 8.8, 29.9, 62.0, 104.7, 159.1, 222.0, 294.5, 380.4, 471.1, 571.7, 686.8, 809.2.
    a. 求出这些数据的最小二乘立方曲线 $y=\beta_0+\beta_1 t+\beta_2 t^2+\beta_3 t^3$.
    b. 利用（a）的结果估计当 $t=4.5$ 秒时飞机的水平速度.

14. 令 $\overline{x}=\dfrac{1}{n}(x_1+\cdots+x_n)$ 和 $\overline{y}=\dfrac{1}{n}(y_1+\cdots+y_n)$，证明：数据 $(x_1,y_1),\cdots,(x_n,y_n)$ 的最小二乘直线必须通过 $(\overline{x},\overline{y})$，也就是证明 $\overline{x}$ 和 $\overline{y}$ 适合线性方程 $\overline{y}=\hat{\beta}_0+\hat{\beta}_1\overline{x}$.（提示：从向量方程 $\boldsymbol{y}=X\hat{\boldsymbol{\beta}}+\boldsymbol{\varepsilon}$ 导出这个方程，将 X 的第一列表示为 $\mathbf{1}$，利用余差向量 $\boldsymbol{\varepsilon}$ 与 X 的列空间正交的事实得到它正交于 $\mathbf{1}$.）

给定数据的最小二乘问题，$(x_1,y_1),\cdots,(x_n,y_n)$，下列缩写符号很有用：

$$\sum x=\sum_{i=1}^{n} x_i \qquad \sum x^2=\sum_{i=1}^{n} x_i^2$$
$$\sum y=\sum_{i=1}^{n} y_i \qquad \sum xy=\sum_{i=1}^{n} x_i y_i$$

最小二乘直线 $y=\hat{\beta}_0+\hat{\beta}_1 x$ 的标准方程可以写成

$$\begin{aligned} n\hat{\beta}_0+\hat{\beta}_1\sum x&=\sum y \\ \hat{\beta}_0\sum x+\hat{\beta}_1\sum x^2&=\sum xy \end{aligned} \qquad (7)$$

15. 对应标准方程（7），导出本节给定的矩阵形式.

16. 利用矩阵的逆求解方程组（7），然后确定许多统计课本中出现的计算 $\hat{\beta}_0$ 和 $\hat{\beta}_1$ 的公式.

17. a. 用新 x 坐标的平均偏差形式，重写例 1 中的数据，设 X 是对应的设计矩阵，说明为什么 X 的列是正交的.
    b. 写出(a)部分中数据的标准方程，并求解得到最小二乘直线 $y=\beta_0+\beta_1 x^*$，此处 $x^*=x-5.5$.

18. 若数据 $(x_1,y_1),\cdots,(x_n,y_n)$ 的 x 坐标是平均偏差形式，所以 $\sum x_i=0$，证明：如果 X 是这种情形下最小二乘直线的设计矩阵，那么 $X^{\mathrm{T}}X$ 是正交矩阵.

习题 19～20 包含具有 2 列或更多列的设计矩阵 X，$\hat{\boldsymbol{\beta}}$ 是 $\boldsymbol{y}=X\boldsymbol{\beta}$ 的最小二乘解，考虑下列数：

（i）$\|X\hat{\boldsymbol{\beta}}\|^2$ ——“回归项”的平方和，记该数为 $SS(R)$.

（ii）$\|\boldsymbol{y}-X\hat{\boldsymbol{\beta}}\|^2$——误差项的平方和，记该数为 $SS(E)$.

（iii）$\|\boldsymbol{y}\|^2$——坐标 y 平方之和的“总和”，记该数为 $SS(T)$.

每一本讨论回归和线性模型 $\boldsymbol{y}=X\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ 的统计课本都引入这些数，尽管术语和记号会有变化，为简单起见，假设 y 的平均值是零，在这种情形下，$SS(T)$ 与被称为 y 值集合的方差成正比.

19. 验证方程 $SS(T)=SS(R)+SS(E)$（提示：利用一个定理并解释为什么定理的假设成立）. 这个方程在统计学中的回归理论和方差分析中非常重要.

20. 证明：$\|X\hat{\boldsymbol{\beta}}\|^2=\hat{\boldsymbol{\beta}}^{\mathrm{T}}X^{\mathrm{T}}\boldsymbol{y}$.（提示：重写方程左边并且利用 $\hat{\boldsymbol{\beta}}$ 满足标准方程的事实.）关于 $SS(R)$ 的这个公式被用在统计学上，从这个结论和习题 19 可得最小二乘误差的标准公式：

$$SS(E)=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}-\hat{\boldsymbol{\beta}}^{\mathrm{T}}X^{\mathrm{T}}\boldsymbol{y}$$

练习题答案

我们必须构造 X 和 $\boldsymbol{\beta}$，使得 $X\boldsymbol{\beta}$ 的第 k 行是对应数据 (x_k, y_k) 的预测 y 值，记为：

$$\beta_0 + \beta_1 x_k + \beta_2 \sin(2\pi x_k/12)$$

显然有

$$X = \begin{bmatrix} 1 & x_1 & \sin(2\pi x_1/12) \\ \vdots & \vdots & \vdots \\ 1 & x_n & \sin(2\pi x_n/12) \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

见图 6-33.

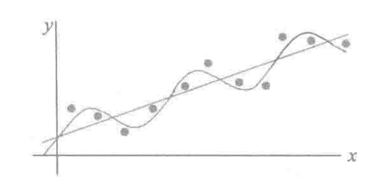


图 6-33　具有季节性波动的销售趋势

6.7　内积空间

长度、距离和正交性的概念在向量空间中有非常重要的应用，对 $\mathbb{R}^n$ 空间，这些概念基于 6.1 节定理 1 中列出的内积性质. 对其他空间，我们需要类似的内积和同样的性质，定理 1 中的结论成为下面定义中的公理.

定义　向量空间 V 上的**内积**是一个函数，对每一对属于 V 的向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$，存在一个实数 $\langle \boldsymbol{u}, \boldsymbol{v} \rangle$ 满足下面公理，对任意属于 V 的 $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ 和所有数 c：

1. $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \boldsymbol{u} \rangle$.
2. $\langle \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{u}, \boldsymbol{w} \rangle + \langle \boldsymbol{v}, \boldsymbol{w} \rangle$.
3. $\langle c\boldsymbol{u}, \boldsymbol{v} \rangle = c\langle \boldsymbol{u}, \boldsymbol{v} \rangle$.
4. $\langle \boldsymbol{u}, \boldsymbol{u} \rangle \geqslant 0$，且 $\langle \boldsymbol{u}, \boldsymbol{u} \rangle = 0$ 的充分必要条件是 $\boldsymbol{u} = \boldsymbol{0}$.

一个赋予上面内积的向量空间称为**内积空间**.

具有标准内积的向量空间 $\mathbb{R}^n$ 是一个内积空间，而且本章几乎所有 $\mathbb{R}^n$ 空间上的讨论都在内积空间上. 本节和 6.8 节的例子给出许多基础应用实例，涉及工程、物理、数学和统计等课程.

例 1　给定两个正整数（例如 4 和 5）及 $\mathbb{R}^2$ 中向量 $\boldsymbol{u} = (u_1, u_2)$ 和 $\boldsymbol{v} = (v_1, v_2)$，规定

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 4u_1v_1 + 5u_2v_2 \tag{1}$$

说明（1）定义了一个内积.

解　当然，公理 1 满足，因为

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 4u_1v_1 + 5u_2v_2 = 4v_1u_1 + 5v_2u_2 = \langle \boldsymbol{v}, \boldsymbol{u} \rangle$$

如果 $\boldsymbol{w} = (w_1, w_2)$，那么

$$\begin{aligned}\langle \boldsymbol{u}+\boldsymbol{v},\boldsymbol{w}\rangle &= 4(u_1+v_1)w_1+5(u_2+v_2)w_2\\ &= 4u_1w_1+5u_2w_2+4v_1w_1+5v_2w_2\\ &= \langle \boldsymbol{u},\boldsymbol{w}\rangle+\langle \boldsymbol{v},\boldsymbol{w}\rangle\end{aligned}$$

这就验证了公理2，对公理3，我们有

$$\begin{aligned}\langle c\boldsymbol{u},\boldsymbol{v}\rangle &= 4(cu_1)v_1+5(cu_2)v_2\\ &= c(4u_1v_1+5u_2v_2)=c\langle \boldsymbol{u},\boldsymbol{v}\rangle\end{aligned}$$

对公理 4，注意 $\langle \boldsymbol{u},\boldsymbol{u}\rangle=4u_1^2+5u_2^2\geqslant 0$ 且 $4u_1^2+5u_2^2=0$，当且仅当 $u_1=u_2=0$ 时成立，即 $\boldsymbol{u}=\boldsymbol{0}$，此外，$\langle \boldsymbol{0},\boldsymbol{0}\rangle=0$，所以（1）定义了 $\mathbb{R}^2$ 上的内积. ■

类似（1）可在 $\mathbb{R}^n$ 上定义内积，他们自然和“带权值的最小二乘”问题联系起来. 此时，权值可赋给内积中和式的各个分量，且这种方式对较重要的分量给予更可靠的权值.

从现在起，当内积空间包含多项式或其他函数时，我们可用悉熟的方式写出函数，而不用黑体表示向量，然而，必须要记住，当它作为向量空间一个元素时每个函数表示的是一个向量.

例 2 设 $t_0,\cdots,t_n$ 是不同的实数，对属于 $\mathbb{P}_n$ 中的 $\boldsymbol{p}$ 和 $\boldsymbol{q}$，定义

$$\langle p,q\rangle=p(t_0)q(t_0)+p(t_1)q(t_1)+\cdots+p(t_n)q(t_n) \tag{2}$$

很容易验证内积公理的 1~3，对公理 4，注意

$$\langle p,p\rangle=[p(t_0)]^2+[p(t_1)]^2+\cdots+[p(t_n)]^2\geqslant 0$$

也有 $\langle \boldsymbol{0},\boldsymbol{0}\rangle=0$.（我们仍用黑体 $\boldsymbol{0}$ 表示零多项式，它是 $\mathbb{P}_n$ 中的零向量.）如果 $\langle p,p\rangle=0$，那么 $\boldsymbol{p}$ 一定在 $n+1$ 个点 $t_0,\cdots,t_n$ 处为零，这时 p 只能是零多项式，因为 p 的次数小于 $n+1$，从而（2）定义了 $\mathbb{P}_n$ 上的一个内积. ■

例3 设 V 属于 $\mathbb{P}_2$，且 V 具有例2中的内积，其中 $t_0=0$，$t_1=\dfrac{1}{2}$ 和 $t_2=1$. 设 $p(t)=12t^2$ 和 $q(t)=2t-1$. 计算 $\langle p,q\rangle$ 和 $\langle q,q\rangle$.

解

$$\begin{aligned}\langle p\cdot q\rangle &= p(0)q(0)+p\left(\frac{1}{2}\right)q\left(\frac{1}{2}\right)+p(1)q(1)\\ &= (0)(-1)+(3)(0)+(12)(1)=12\\ \langle q\cdot q\rangle &= [q(0)]^2+\left[q\left(\frac{1}{2}\right)\right]^2+[q(1)]^2\\ &= (-1)^2+(0)^2+(1)^2=2\end{aligned}$$

■

长度、距离和正交性

设 V 是一个内积空间，其内积记作 $\langle \boldsymbol{u},\boldsymbol{v}\rangle$，像 $\mathbb{R}^n$ 空间一样，我们定义一个向量的**长度**或**范数**是数

$$\|\boldsymbol{v}\|=\sqrt{\langle \boldsymbol{v},\boldsymbol{v}\rangle}$$

即 $\|\boldsymbol{v}\|^2=\langle \boldsymbol{v},\boldsymbol{v}\rangle$（这个定义有意义，因为 $\langle \boldsymbol{v},\boldsymbol{v}\rangle\geqslant 0$，但这个意义并不是说 $\langle \boldsymbol{v},\boldsymbol{v}\rangle$ 是一个“平方之和”，因为 $\boldsymbol{v}$ 不必是 $\mathbb{R}^n$ 中的元素）.

一个**单位向量**是长度为 1 的向量，**向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 之间的距离**是 $\|\boldsymbol{u}-\boldsymbol{v}\|$，向量 $\boldsymbol{u}$ 和向量 $\boldsymbol{v}$ **正交**，如果 $\langle \boldsymbol{u},\boldsymbol{v}\rangle = 0$ 成立.

例 4 若 $\mathbb{P}_2$ 具有例 3 中的形式（2）的内积，计算向量 $p(t)=12t^2$ 和 $q(t)=2t-1$ 的长度.

解

$$\|p\|^2 = \langle p,p\rangle = [p(0)]^2 + \left[p\left(\frac{1}{2}\right)\right]^2 + [p(1)]^2 = 0+[3]^2+[12]^2 = 153$$

$$\|p\| = \sqrt{153}$$

在例 3 中，我们知道 $\langle q,q\rangle = 2$，因此 $\|q\| = \sqrt{2}$. ■

格拉姆-施密特方法

内积空间中有限维子空间的正交基的存在性，可由格拉姆-施密特方法确定，像 $\mathbb{R}^n$ 空间一样. 应用中经常出现的一些正交基可用这个方法构造.

一个向量在一个具有正交基的子空间 W 上的正交投影，可像平常一样构造. 投影不依靠正交基的选取，并且它们有正交分解定理和最佳逼近定理所描述的性质.

例 5 若 V 是具有例 2 中内积的 $\mathbb{P}_4$，包含多项式在−2，−1，0，1 和 2 处的值，且 $\mathbb{P}_2$ 作为 V 的一个子空间，应用格拉姆-施密特方法于多项式 1，t 和 t^2，构造 $\mathbb{P}_2$ 的一个正交基.

解 内积仅依赖于多项式在−2，−1，0，1 和 2 的值，所以我们列出每个多项式作为 $\mathbb{R}^5$ 空间的值，写在多项式的下面：[⊖]

$$\begin{array}{cccc} \text{多项式：} & 1 & t & t^2 \\ \text{向量值：} & \begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix} & \begin{bmatrix}-2\\-1\\0\\1\\2\end{bmatrix} & \begin{bmatrix}4\\1\\0\\1\\4\end{bmatrix} \end{array}$$

V 中两个多项式的内积等于它们对应向量在 $\mathbb{R}^5$ 空间的标准内积. 注意 t 与常数函数 1 正交，所以取 $p_0(t)=1$ 和 $p_1(t)=t$，对 p_2，利用 $\mathbb{R}^5$ 中的向量，计算 t^2 在 $\text{Span}\{p_0,p_1\}$ 上的投影.

$$\langle t^2,p_0\rangle = \langle t^2,1\rangle = 4+1+0+1+4 = 10$$

$$\langle p_0,p_0\rangle = 5$$

$$\langle t^2,p_1\rangle = \langle t^2,t\rangle = -8+(-1)+0+1+8 = 0$$

t^2 在 $\text{Span}\{1,t\}$ 上的正交投影是 $\frac{10}{5}p_0 + 0\cdot p_1$，这样

$$p_2(t) = t^2 - 2p_0(t) = t^2 - 2$$

V 的子空间 $\mathbb{P}_2$ 的一个正交基是：

⊖ $\mathbb{P}_4$ 中的每个多项式被其在 5 个 t 值−2，−1，0，1 和 2 上的取值惟一确定. 事实上，p 和它的向量值的对应关系是一种同构关系亦即是保持线性组合关系的一对一映上到 $\mathbb{R}^5$ 的映射.

多项式： p_0 p_1 p_2

$$\text{向量值：}\begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}\quad\begin{bmatrix}-2\\-1\\0\\1\\2\end{bmatrix}\quad\begin{bmatrix}2\\-1\\-2\\-1\\2\end{bmatrix}\qquad(3)$$

■

内积空间的最佳逼近

应用数学中最常见的问题，包括元素是函数的向量空间，主要是在V的特定子空间W中，选取函数g来逼近V中的函数f．对f的“逼近”程度依赖$\|f-g\|$定义的方式，我们仅考虑f和g的距离用内积定义的情形，在此情形下，f由W中函数的最佳逼近是指f在子空间W上的正交投影.

例6 设V是$\mathbb{P}_4$具有例5中定义的内积，p_0, p_1和p_2是例5中的子空间$\mathbb{P}_2$的正交基，求出$\mathbb{P}_2$中的多项式对$p(t)=5-\frac{1}{2}t^4$的最佳逼近.

解 p_0, p_1和p_2对应t为−2，−1，0，1和2的值以$\mathbb{R}^5$中向量的形式在上面（3）中已经给出，p的相应值是−3，9/2，5, 9/2和−3，我们计算：

$$\langle p, p_0\rangle = 8, \quad \langle p, p_1\rangle = 0, \quad \langle p, p_2\rangle = -31$$
$$\langle p_0, p_0\rangle = 5, \quad \langle p_2, p_2\rangle = 14$$

所以，$\mathbb{P}_2$中的多项式对V中p的最佳逼近是

$$\hat{p} = \text{proj}_{\mathbb{P}_2} p = \frac{\langle p, p_0\rangle}{\langle p_0, p_0\rangle} p_0 + \frac{\langle p, p_1\rangle}{\langle p_1, p_1\rangle} p_1 + \frac{\langle p, p_2\rangle}{\langle p_2, p_2\rangle} p_2$$
$$= \frac{8}{5} p_0 + \frac{-31}{14} p_2 = \frac{8}{5} - \frac{31}{14}(t^2-2)$$

当多项式之间的距离仅用t为−2，−1，0，1和2时的值来度量时，这是$\mathbb{P}_2$的所有多项式中离p最接近的多项式，见图6-34.

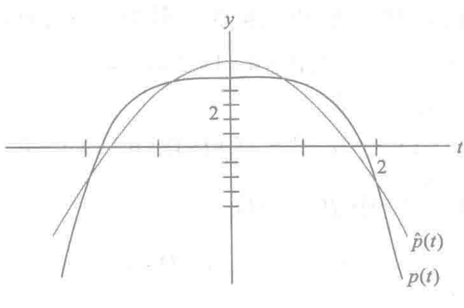


图 6-34

■

例 5 和例 6 中的多项式 p_0, p_1 和 p_2 属于一类多项式，在统计学上称为正交多项式.[⊖] 正交性是指例 2 中描述的内积类型.

两个不等式

给定内积空间中的向量 $\boldsymbol{v}$ 和给定的有限维子空间 W . 我们将勾股定理应用到 $\boldsymbol{v}$ 相对 W 的正交分解中，可以得到：

$$\|\boldsymbol{v}\|^2 = \|\text{proj}_W \boldsymbol{v}\|^2 + \|\boldsymbol{v} - \text{proj}_W \boldsymbol{v}\|^2$$

见图 6-35，特别地，这表明 $\boldsymbol{v}$ 到 W 上投影的范数，不超过 $\boldsymbol{v}$ 自己的范数，这个简单事实可推出下面重要的不等式.

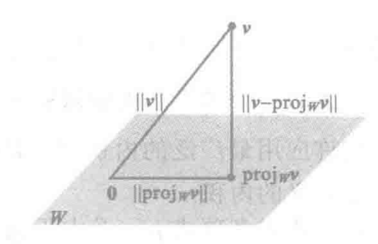


图 6-35　直角三角形的斜边是最长边

定理 16　（柯西-施瓦茨不等式）

对空间 V 中任意向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ ，有

$$|\langle \boldsymbol{u}, \boldsymbol{v}\rangle| \leqslant \|\boldsymbol{u}\| \ \|\boldsymbol{v}\| \tag{4}$$

证　如果 $\boldsymbol{u} = \boldsymbol{0}$ ，方程（4）的两边都是零，且这种情形下（4）式成立（见练习题 1）. 如果 $\boldsymbol{u} \neq \boldsymbol{0}$ ，若 W 是 $\boldsymbol{u}$ 生成的子空间，注意 $\|c\boldsymbol{u}\| = |c| \ \|\boldsymbol{u}\|$ 对任何数 c 都成立. 因此

$$\|\text{proj}_W \boldsymbol{v}\| = \left\| \frac{\langle \boldsymbol{v}, \boldsymbol{u}\rangle}{\langle \boldsymbol{u}, \boldsymbol{u}\rangle} \boldsymbol{u} \right\| = \frac{|\langle \boldsymbol{v}, \boldsymbol{u}\rangle|}{|\langle \boldsymbol{u}, \boldsymbol{u}\rangle|} \|\boldsymbol{u}\| = \frac{|\langle \boldsymbol{v}, \boldsymbol{u}\rangle|}{\|\boldsymbol{u}\|^2} \|\boldsymbol{u}\| = \frac{|\langle \boldsymbol{u}, \boldsymbol{v}\rangle|}{\|\boldsymbol{u}\|}$$

由于 $\|\text{proj}_W \boldsymbol{v}\| \leqslant \|\boldsymbol{v}\|$ ，我们得到 $\dfrac{|\langle \boldsymbol{u}, \boldsymbol{v}\rangle|}{\|\boldsymbol{u}\|} \leqslant \|\boldsymbol{v}\|$ ，即给出（4）式. ■

柯西-施瓦茨不等式在很多数学分支都很有用，习题中有几个简单应用. 这里我们主要利用它证明另一个包含向量范数的基本不等式. 见图 6-36.

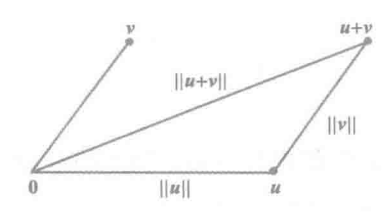


图 6-36　三角形边的长度

定理 17　（三角不等式）

对属于 V 的所有向量 $\boldsymbol{u}, \boldsymbol{v}$ ，有

$$\|\boldsymbol{u} + \boldsymbol{v}\| \leqslant \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$$

⊖ 见 *Statistics and Experimental Design in Engineering and the Physical Sciences*, 2nd ed, Norman L. Johnson and Fred C. Leone （New York: John Wiley & Sons, 1977），那里表格列出的"正交多项式"，是指多项式在−2，−1，0，1，2 处的值.

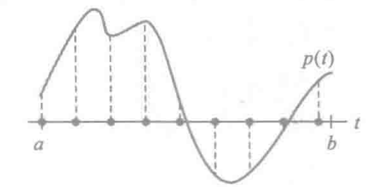
证
$$\begin{aligned}\|u+v\|^2 &= \langle u+v, u+v\rangle = \langle u,u\rangle + 2\langle u,v\rangle + \langle v,v\rangle \\ &\leqslant \|u\|^2 + 2|\langle u,v\rangle| + \|v\|^2 \\ &\leqslant \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \qquad \text{柯西-施瓦茨不等式} \\ &= (\|u\|+\|v\|)^2\end{aligned}$$

两边开方后，立刻得到三角不等式. ■

$C[a,b]$ 上的一个内积（需要微积分知识）

也许应用最广泛的内积空间是区间 $a \leqslant t \leqslant b$ 上所有连续函数构成的向量空间 $C[a,b]$，且具有下面定义的内积.

首先考虑多项式 p，及大于或等于 p 的阶数 n 的任何整数，则有 p 属于 $\mathbb{P}_n$，我们可以利用例 2 中的内积计算“长度”，注意 p 包含 $[a,b]$ 中 $(n+1)$ 个点. 然而，这种 p 的长度仅保留这 $(n+1)$ 个点的特性，一般地，对 p 属于 $\mathbb{P}_n$ 的所有 n，我们可以利用更大的数 n，更多的求值点计算对应的内积. 见图 6-37.



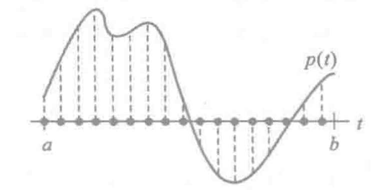
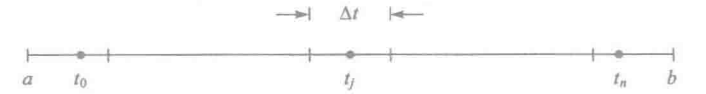


图 6-37 利用 $[a,b]$ 内不同数目的求值点计算 $\|p\|^2$

我们用 $(n+1)$ 个长度为 $\Delta t = (b-a)/(n+1)$ 的子区间分割 $[a,b]$，并且使 $t_0,\cdots,t_n$ 是这些子区间中的任意点.



如果 n 很大，由 $t_0,\cdots,t_n$ 确定的关于 $\mathbb{P}_n$ 的内积将趋向较大的数 $\langle p,p\rangle$，所以需要重新度量，并将内积除以 $(n+1)$，注意到 $1/(n+1) = \Delta t/(b-a)$，定义：

$$\langle p,q\rangle = \frac{1}{n+1}\sum_{j=0}^{n} p(t_j)q(t_j) = \frac{1}{b-a}\left[\sum_{j=0}^{n} p(t_j)q(t_j)\Delta t\right]$$

现在，让 n 无限制增加，由于多项式 p,q 是连续函数，括号内的表达式是一个黎曼和且趋向一个定积分，考虑 $p(t)q(t)$ 在 $[a,b]$ 上的平均值：

$$\frac{1}{b-a}\int_a^b p(t)q(t)\mathrm{d}t$$

这个数值对任意阶多项式都有定义（事实上是对所有连续函数），且它具有像下面例题所说的全部内积性质，前面的因子 $1/(b-a)$ 不是必需的，为简化下面的计算常省略.

例 7 对 $C[a,b]$ 中的 f,g，取

$$\langle f,g\rangle = \int_a^b f(t)g(t)\mathrm{d}t \tag{5}$$

这证明（5）定义了 $C[a,b]$ 上的内积.

解 内积公理 1~3 可由积分的基本性质得出，对公理 4，注意到

$$\langle f,f\rangle = \int_a^b [f(t)]^2\,\mathrm{d}t \geqslant 0$$

函数 $[f(t)]^2$ 在 $[a,b]$ 上连续且非负，如果 $[f(t)]^2$ 的定积分为零，那么由高等微积分的定理可知，$[f(t)]^2$ 在 $[a,b]$ 上必须恒等于零. 从而 f 是一个零函数，$\langle f,f\rangle=0$ 意味着 f 是 $[a,b]$ 上的零函数，也就是（5）定义了一个 $[a,b]$ 上的内积. ■

例 8 设 V 表示内积用例 7 定义的内积空间 $C[0,1]$，W 是多项式 $p_1(t)=1$，$p_2(t)=2t-1$ 和 $p_3(t)=12t^2$ 所生成的子空间，利用格拉姆-施密特方法，求 W 的一个正交基.

解 取 $q_1=p_1$，并且计算

$$\langle p_2,q_1\rangle = \int_0^1 (2t-1)(1)\mathrm{d}t = (t^2-t)\Big|_0^1 = 0$$

这样 p_2 已经与 q_1 正交，所以可取 $q_2=p_2$. 对 p_3 在 $W_2=\mathrm{Span}\{q_1,q_2\}$ 的投影，我们计算

$$\langle p_3,q_1\rangle = \int_0^1 12t^2\cdot 1\mathrm{d}t = 4t^3\Big|_0^1 = 4$$

$$\langle q_1,q_1\rangle = \int_0^1 1.1\mathrm{d}t = t\Big|_0^1 = 1$$

$$\langle p_3,q_2\rangle = \int_0^1 12t^2(2t-1)\mathrm{d}t = \int_0^1 (24t^3-12t^2)\mathrm{d}t = 2$$

$$\langle q_2,q_2\rangle = \int_0^1 (2t-1)^2\mathrm{d}t = \frac{1}{6}(2t-1)^3\Big|_0^1 = \frac{1}{3}$$

那么

$$\mathrm{proj}_{W_2} p_3 = \frac{\langle p_3,q_1\rangle}{\langle q_1,q_1\rangle}q_1 + \frac{\langle p_3,q_2\rangle}{\langle q_2,q_2\rangle}q_2 = \frac{4}{1}q_1 + \frac{2}{1/3}q_2 = 4q_1+6q_2$$

$$q_3 = p_3 - \mathrm{proj}_{W_2} p_3 = p_3 - 4q_1 - 6q_2$$

作为一个函数，$q_3(t)=12t^2-4-6(2t-1)=12t^2-12t+2$，子空间 W 的正交基是 $\{q_1,q_2,q_3\}$. ■

练习题

利用内积公理验证下列论断.

1. $\langle \boldsymbol{v},\boldsymbol{0}\rangle=\langle \boldsymbol{0},\boldsymbol{v}\rangle=0$
2. $\langle \boldsymbol{u},\boldsymbol{v}+\boldsymbol{w}\rangle=\langle \boldsymbol{u},\boldsymbol{v}\rangle+\langle \boldsymbol{u},\boldsymbol{w}\rangle$

习题 6.7

1. 在 $\mathbb{R}^2$ 空间中取例 1 定义的内积，$\boldsymbol{x}=(1,1)$，$\boldsymbol{y}=(5,-1)$.
   a. 计算 $\|\boldsymbol{x}\|$, $\|\boldsymbol{y}\|$ 和 $|\langle \boldsymbol{x},\boldsymbol{y}\rangle|^2$.
   b. 描述所有与 $\boldsymbol{y}$ 正交的向量 (z_1,z_2).
2. 在 $\mathbb{R}^2$ 空间中取例 1 定义的内积，说明柯西-施瓦茨不等式，对 $\boldsymbol{x}=(3,-2)$ 和 $\boldsymbol{y}=(-2,1)$ 成立.（建议：研究 $|\langle \boldsymbol{x},\boldsymbol{y}\rangle|^2$.）

习题 3~8 中的多项式属于 $\mathbb{P}_2$ 且计算内积时 t 取值为−1，0 和 1.（见例 2.）

3. 计算 $\langle p,q\rangle$，此处 $p(t)=4+t, q(t)=5-4t^2$.
4. 计算 $\langle p,q\rangle$，此处 $p(t)=3t-t^2, q(t)=3+2t^2$.
5. 计算 $\|p\|$ 和 $\|q\|$，其中 p ,q 如习题 3.

6. 计算$\|p\|$和$\|q\|$，其中 p ,q 如习题 4.
7. 计算 q 在 p 所生成子空间上的正交投影，其中 p ,q 如习题 3.
8. 计算 q 在 p 所生成子空间上的正交投影，其中 p ,q 如习题 4.
9. 设 $\mathbb{P}_3$ 计算内积时 t 取值为−3，−1，1 和 3，取 $p_0(t)=1$, $p_1(t)=t$ 和 $p_2(t)=t^2$.
   a. 计算 p_2 在 p_0 和 p_1 生成的子空间上的正交投影.
   b. 求一个与 p_0 和 p_1 都正交的多项式 q ,使得 $\{p_0,p_1,q\}$ 是 $\text{Span}\{p_0,p_1,p_2\}$ 的一个正交基，重新度量多项式 q ，使得它在 $(-3,-1,1,3)$ 处的向量值是 $(1,-1,-1,1)$.
10. 若 $\mathbb{P}_3$ 具有习题 9 中定义的内积，且多项式 p_0,p_1,q 也如习题 9 定义，求 $\text{Span}\{p_0,p_1,q\}$ 中的多项式对 $p(t)=t^3$ 的最佳逼近.
11. 若 p_0,p_1和p_2 是例 5 中给出的正交多项式，且计算 $\mathbb{P}_4$ 的内积时 t 取值为 $-2,-1,0,1$ 和 2. 求 t^3 在 $\text{Span}\{p_0,p_1,p_2\}$ 上的正交投影.
12. 求一个多项式 p_3，使得 $\{p_0,p_1,p_2,p_3\}$（见习题 11）是 $\mathbb{P}_4$ 中子空间 $\mathbb{P}_3$ 的正交基，重新度量多项式 p_3，使得它的向量值是 $(-1,2,0,-2,1)$.
13. 设 A 是任何 $n\times n$ 可逆矩阵，说明对 $\mathbb{R}^n$ 中 $\boldsymbol{u}$ 和 $\boldsymbol{v}$，公式 $\langle \boldsymbol{u},\boldsymbol{v}\rangle=(A\boldsymbol{u})\cdot(A\boldsymbol{v})=(A\boldsymbol{u})^{\mathrm{T}}\cdot(A\boldsymbol{v})$ 定义了一个 $\mathbb{R}^n$ 上的内积.
14. 若 T 是从向量空间 V 到 $\mathbb{R}^n$ 的一对一线性变换，证明对 V 中的 $\boldsymbol{u}$ 和 $\boldsymbol{v}$，公式 $\langle \boldsymbol{u},\boldsymbol{v}\rangle=T(\boldsymbol{u})\cdot T(\boldsymbol{v})$ 定义了 V 上的一个内积.

利用本节内积公理和其他结果去验证习题 15~18 的命题.

15. 验证 $\langle \boldsymbol{u},c\boldsymbol{v}\rangle=c\langle \boldsymbol{u},\boldsymbol{v}\rangle$ 对所有数 c 都成立.
16. 如果 $\{\boldsymbol{u},\boldsymbol{v}\}$ 是 V 中的单位正交集，那么 $\|\boldsymbol{u}-\boldsymbol{v}\|=\sqrt{2}$.
17. $\langle \boldsymbol{u},\boldsymbol{v}\rangle=\frac{1}{4}\|\boldsymbol{u}+\boldsymbol{v}\|^2-\frac{1}{4}\|\boldsymbol{u}-\boldsymbol{v}\|^2$.
18. $\|\boldsymbol{u}+\boldsymbol{v}\|^2+\|\boldsymbol{u}-\boldsymbol{v}\|^2=2\|\boldsymbol{u}\|^2+2\|\boldsymbol{v}\|^2$.
19. 给定 $a\geqslant 0$ 和 $b\geqslant 0$，设 $\boldsymbol{u}=\begin{bmatrix}\sqrt{a}\\\sqrt{b}\end{bmatrix}$ 和 $\boldsymbol{v}=\begin{bmatrix}\sqrt{b}\\\sqrt{a}\end{bmatrix}$，利用柯西-施瓦茨不等式比较几何平均值 $\sqrt{ab}$ 和算术平均值 $(a+b)/2$.
20. 设 $\boldsymbol{u}=\begin{bmatrix}a\\b\end{bmatrix}$ 和 $\boldsymbol{v}=\begin{bmatrix}1\\1\end{bmatrix}$，利用柯西-施瓦茨不等式证明 $\left(\frac{a+b}{2}\right)^2\leqslant\frac{a^2+b^2}{2}$.

习题 21~24 的空间指的是 $V=C[0,1]$，其内积如例 7 所示用一个积分给出.

21. 计算 $\langle f,g\rangle$，此处 $f(t)=1-3t^2$， $g(t)=t-t^3$.
22. 计算 $\langle f,g\rangle$，此处 $f(t)=5t-3$， $g(t)=t^3-t^2$.
23. 计算习题 21 中 f 的 $\|f\|$.
24. 计算习题 22 中 g 的 $\|g\|$.
25. 设 V 是具有例 7 中定义的内积，空间 $C[-1,1]$，求由多项式 1，t 和 t^2 所生成的子空间的正交基. 这个基中的多项式称为*勒让德多项式*.
26. 设 V 是具有例 7 中定义的内积，空间 $C[-2,2]$，求由多项式 1，t 和 t^2 所生成的子空间的一个正交基.
27. [M]设 $\mathbb{P}_4$ 具有例 5 中定义的内积，且 p_0,p_1,p_2 是该例中的正交多项式，利用矩阵程序，将格拉姆-施密特方法用于集合 $\{p_0,p_1,p_2,t^3,t^4\}$，构造 $\mathbb{P}_4$ 的一个正交基.
28. [M]设 V 是具有例 7 中定义的内积，空间 $C[0,2\pi]$，利用格拉姆-施密特方法构造由 $\{1,\cos t,\cos^2 t,\cos^3 t\}$ 所生成的子空间的一个正交基，利用矩阵程序或计算程序来计算相应的定积分.

练习题答案

1. 由公理 1，$\langle \boldsymbol{v},\boldsymbol{0}\rangle=\langle \boldsymbol{0},\boldsymbol{v}\rangle$，那么 $\langle \boldsymbol{0},\boldsymbol{v}\rangle=\langle 0\boldsymbol{v},\boldsymbol{v}\rangle=0\langle \boldsymbol{v},\boldsymbol{v}\rangle$，由公理 3，得到 $\langle \boldsymbol{0},\boldsymbol{v}\rangle=0$.
2. 由公理 1 和 2，再由公理 1，得 $\langle \boldsymbol{u},\boldsymbol{v}+\boldsymbol{w}\rangle=\langle \boldsymbol{v}+\boldsymbol{w},\boldsymbol{u}\rangle=\langle \boldsymbol{v},\boldsymbol{u}\rangle+\langle \boldsymbol{w},\boldsymbol{u}\rangle=\langle \boldsymbol{u},\boldsymbol{v}\rangle+\langle \boldsymbol{u},\boldsymbol{w}\rangle$.

6.8 内积空间的应用

本节的例题说明 6.7 节定义的内积空间，如何应用在实际问题中，第一个例子是本章介绍中的例子，它是调整北美地质数据问题中，出现的巨型最小二乘问题.

加权最小二乘法

设向量 $\boldsymbol{y}$ 的 n 次观测值为 $y_1,\cdots,y_n$，且假设我们希望，用属于 $\mathbb{R}^n$ 的特定子空间中一个向量 $\hat{\boldsymbol{y}}$ 逼近 $\boldsymbol{y}$（在 6.5 节，$\hat{\boldsymbol{y}}$ 被写成 $A\boldsymbol{x}$，所以 $\hat{\boldsymbol{y}}$ 属于 A 的列空间）. 记 $\hat{\boldsymbol{y}}$ 的分量为 $\hat{y}_1,\cdots,\hat{y}_n$，那么误差的平方和或 $SS(E)$，用 $\hat{\boldsymbol{y}}$ 逼近 $\boldsymbol{y}$ 后，得到

$$SS(E)=(y_1-\hat{y}_1)^2+\cdots+(y_n-\hat{y}_n)^2 \tag{1}$$

利用 $\mathbb{R}^n$ 的标准长度的写法，上式可简记为 $\|\boldsymbol{y}-\hat{\boldsymbol{y}}\|^2$.

现在，假设测量时 y 的各个分量的可靠性不同（这是北美地质资料的一个特点，因为测量的数据是 140 年前的，作为另一个例子，y 的分量的计算来自各种样本的测量和不同样本的大小），那么可靠性就变成（1）式中平方误差的适当权值，较可靠的测量应赋予更重要的作用.[㊀] 如果权值记为 $w_1^2,\cdots,w_n^2$，那么加权的误差平方和是

$$\text{加权 } SS(E)=w_1^2(y-\hat{y}_1)^2+\cdots+w_n^2(y_n-\hat{y}_n)^2 \tag{2}$$

这是 $(\boldsymbol{y}-\hat{\boldsymbol{y}})$ 长度的平方，此处的长度类似 6.7 节例 1 中定义的内积，记为

$$\langle \boldsymbol{x},\boldsymbol{y}\rangle = w_1^2x_1y_1+\cdots+w_n^2x_ny_n$$

有时，可以非常方便地将这种加权最小二乘问题变换为等价的普通最小二乘问题，设 W 是对角线上是非负数 $w_1,\cdots,w_n$ 的对角矩阵，可得：

$$W\boldsymbol{y}=\begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & \\ \vdots & & \ddots & \vdots \\ 0 & & \cdots & w_n \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}=\begin{bmatrix} w_1y_1 \\ w_2y_2 \\ \vdots \\ w_ny_n \end{bmatrix}$$

$W\hat{\boldsymbol{y}}$ 有类似的表达式，观察到（2）式的第 j 项可写成

$$w_j^2(y_j-\hat{y}_j)^2=(w_jy_j-w_j\hat{y}_j)^2$$

可见（2）式中加权的 $SS(E)$ 就是 $\mathbb{R}^n$ 中 $W\boldsymbol{y}-W\hat{\boldsymbol{y}}$ 普通长度的平方，它可以写成 $\|W\boldsymbol{y}-W\hat{\boldsymbol{y}}\|^2$.

现在假设向量 $\hat{\boldsymbol{y}}$ 的逼近是由矩阵 A 的列构成的，我们寻找一个 $\hat{\boldsymbol{x}}$，使得 $A\hat{\boldsymbol{x}}=\hat{\boldsymbol{y}}$ 尽可能接近 $\boldsymbol{y}$. 然而，逼近的度量是含权的误差：

$$\|W\boldsymbol{y}-W\hat{\boldsymbol{y}}\|^2=\|W\boldsymbol{y}-WA\hat{\boldsymbol{x}}\|^2$$

这样 $\hat{\boldsymbol{x}}$ 是方程

$$WA\boldsymbol{x}=W\boldsymbol{y}$$

㊀ 有统计学知识的读者注意：假若 y_i 的测量误差是独立随机变量，且均值为零，方差分别为 $\sigma_1^2,\cdots,\sigma_n^2$，（2）中适当的权值是 $w_i^2=1/\sigma_i^2$，较大的方差误差，对应较小的权值.

的普通最小二乘解，此最小二乘问题的标准方程是：

$$(WA)^{\mathrm{T}} WA\boldsymbol{x} = (WA)^{\mathrm{T}} W\boldsymbol{y}$$

例 1 求最小二乘直线 $y = \beta_0 + \beta_1 x$，最佳拟合数据为 $(-2,3),(-1,5),(0,5),(1,4),(2,3)$，假设后面两组数据中，$y$ 值测量的误差比其余数据的误差大，这些数据的权值只有其余数据权值的一半.

解 如 6.6 节所示，写出矩阵 A 对应的 X，向量 $\boldsymbol{x}$ 对应 $\boldsymbol{\beta}$，我们得到：

$$X = \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} 3 \\ 5 \\ 5 \\ 4 \\ 3 \end{bmatrix}$$

对权矩阵，选取 W 的对角线值为 2，2，2，1 和 1．对 X 的行和 $\boldsymbol{y}$ 分别左乘 W，得到：

$$WX = \begin{bmatrix} 2 & -4 \\ 2 & -2 \\ 2 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad W\boldsymbol{y} = \begin{bmatrix} 6 \\ 10 \\ 10 \\ 4 \\ 3 \end{bmatrix}$$

对于标准方程，计算

$$(WX)^{\mathrm{T}} WX = \begin{bmatrix} 14 & -9 \\ -9 & 25 \end{bmatrix} \text{和} (WX)^{\mathrm{T}} W\boldsymbol{y} = \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

并且求解

$$\begin{bmatrix} 14 & -9 \\ -9 & 25 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

标准方程的解（精确到 2 位有效数字）是 $\beta_0 = 4.3$ 和 $\beta_1 = 0.20$，期望的直线是 $y = 4.3 + 0.20x$．相反，这些数据的普通最小二乘直线是 $y = 4.0 - 0.10x$．两条直线都显示在图 6-38 中.

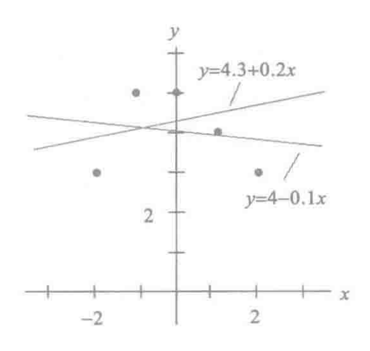


图 6-38 带权值和普通的最小二乘直线 ■

数据趋势分析

若特定函数 f 仅知道在点 $t_0,\cdots,t_n$ 处的值（也许是近似值），如果数据 $f(t_0),\cdots,f(t_n)$ 有一个线性趋势，那么我们期望用形如 $\beta_0+\beta_1 t$ 的函数得到 f 的近似值. 如果有一个“二次趋势”，那么我们会尝试用形如 $\beta_0+\beta_1 t+\beta_2 t^2$ 的函数. 这就是不同的观点下，在 6.6 节已讨论过的函数.

在某些统计问题中，将线性趋势从二次趋势（也许三次或高阶趋势）中分离出来非常重要. 例如，工程师正在分析新车的性能，而 $f(t)$ 表示 t 时刻汽车和一些参照点的距离，如果汽车以常速连续行驶，那么 $f(t)$ 的图像应该是直线且斜率表示速度，如果突然踩下油门，$f(t)$ 的图像将改变为包含二次项和三次项（由于加速的原因）. 又例如，当分析一辆汽车超过另一辆汽车的能力时，工程师就会希望将二次或三次项从一次项中分离出来.

如果一个函数由形如 $y=\beta_0+\beta_1 t+\beta_2 t^2$ 的函数来逼近，系数 β_2 也许不能给出期望的二次趋向数据，原因是在统计学意义下，它和其他 β_i 相关，为进行所谓数据的**趋势分析**，类似 6.7 节的例 2，我们引入空间 $\mathbb{P}_n$ 上的内积，对属于 $\mathbb{P}_n$ 的 p,q，定义：

$$\langle p,q\rangle = p(t_0)q(t_0)+\cdots+p(t_n)q(t_n)$$

实际上，统计学家很少需要考虑阶数高于三次或四次的趋势. 所以，假设 p_0,p_1,p_2,p_3 表示 $\mathbb{P}_n$ 的子空间 $\mathbb{P}_3$ 的正交基，它可以将多项式 $1,t,t^2$ 和 t^3 通过格拉姆-施密特正交化方法得到. 利用第 2 章补充习题 11，存在一个属于 $\mathbb{P}_n$ 的多项式 g，它在 $t_0,\cdots,t_n$ 处的值与未知函数 f 处的值一致. 若 $\hat{g}$ 是 g 在 $\mathbb{P}_3$ 上的正交投影（相对于给定的内积），则

$$\hat{g}=c_0p_0+c_1p_1+c_2p_2+c_3p_3$$

那么 $\hat{g}$ 称为数据的立方**趋势函数**，$c_0,\cdots,c_3$ 称为数据的**趋势系数**. 其中系数 c_1 表示线性趋势，c_2 表示二次趋势，c_3 表示立方趋势. 结果是如果数据具有某些性质，这些系数相互独立.

由于 $p_0,\cdots,p_3$ 是正交的，（注意 $c_i=\langle g,p_i\rangle/\langle p_i,p_i\rangle$.）. 趋势系数可逐次计算且相互独立. 如果我们仅需要二次趋势，我们可以忽略 p_3和c_3. 例如，如果我们需要确定四次趋势，我们仅需要计算 $\langle g,p_4\rangle/\langle p_4,p_4\rangle$，找到一个与 $\mathbb{P}_3$ 正交且属于 p_4 的多项式 p_4（通过格拉姆-施密特方法）.

例 2 最简单且最重要的趋势分析是点 $t_0,\cdots,t_n$ 被调整后，它们均匀分布且总和为零，用二次函数拟合数据 $(-2,3),(-1,5),(0,5),(1,4),(2,3)$.

解 对 6.7 节中例 5 的正交多项式的 t 坐标重新度量，我们可得

$$\begin{array}{lcccc}
\text{多项式：} & p_0 & p_1 & p_2 & \text{数据：}\ g \\
\text{向量值：} & \begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}, & \begin{bmatrix}-2\\-1\\0\\1\\2\end{bmatrix}, & \begin{bmatrix}2\\-1\\-2\\-1\\2\end{bmatrix}, & \begin{bmatrix}3\\5\\5\\4\\3\end{bmatrix}
\end{array}$$

计算仅包含这些向量，没有涉及特别的正交多项式公式，在多项式空间 $\mathbb{P}_2$ 中，数据的最佳拟合是下面给出的最佳投影.

$$\hat{p} = \frac{\langle g, p_0 \rangle}{\langle p_0, p_0 \rangle} p_0 + \frac{\langle g, p_1 \rangle}{\langle p_1, p_1 \rangle} p_1 + \frac{\langle g, p_2 \rangle}{\langle p_2, p_2 \rangle} p_2$$
$$= \frac{20}{5} p_0 - \frac{1}{10} p_1 - \frac{7}{14} p_2$$

且

$$\hat{p}(t) = 4 - 0.1t - 0.5(t^2 - 2) \tag{3}$$

由于 p_2 的系数不是足够小，一个合理的结果是趋势至少是二次. 这个结论可从图 6-39 得到验证.

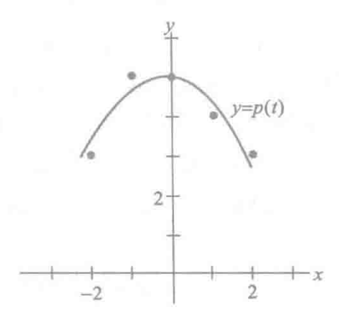


图 6-39　用二次趋势函数逼近 ■

傅里叶级数（需要微积分知识）

连续函数常用正弦和余弦函数的线性组合来逼近，例如，一个连续函数可以表示一个声波，某类电信号或力学振动系统的运动等.

为简单起见，我们考虑 $0 \leqslant t \leqslant 2\pi$ 上的函数，结果是任何 $C[0,2\pi]$ 上的函数可以由下列形式的函数任意逼近：

$$\frac{a_0}{2} + a_1 \cos t + \cdots + a_n \cos nt + b_1 \sin t + \cdots + b_n \sin nt \tag{4}$$

如果自然数 n 足够大.（4）中的函数称为三角多项式，如果 a_n 和 b_n 不同时为零，多项式称为是 n 阶的，三角多项式和 $C[0,2\pi]$ 上的其他函数之间的联系依赖下列事实. 对任何 $n \geqslant 1$，集合

$$\{1, \cos t, \cos 2t, \cdots, \cos nt, \sin t, \sin 2t, \cdots, \sin nt\} \tag{5}$$

对下面定义的内积是正交的：

$$\langle f, g \rangle = \int_0^{2\pi} f(t)g(t)\mathrm{d}t \tag{6}$$

这个正交性可从下面的例题和习题 5 和习题 6 得到验证.

例 3　空间 $C[0,2\pi]$ 具有形如(6)的内积，并且 m 和 n 是不相等的正整数. 证明 $\cos mt$ 和 $\cos nt$ 正交.

解　我们利用三角恒等式，如果 $m \neq n$，有

$$\langle \cos mt, \cos nt \rangle = \int_0^{2\pi} \cos mt \cos nt \mathrm{d}t$$
$$= \frac{1}{2} \int_0^{2\pi} [\cos(mt + nt) + \cos(mt - nt)] \mathrm{d}t$$

$$=\frac{1}{2}\left[\frac{\sin(mt+nt)}{m+n}+\frac{\sin(mt-nt)}{m-n}\right]\Bigg|_0^{2\pi}$$
$$=0$$

■

设 W 是 $C[0,2\pi]$ 中的子空间且由（5）中的函数所生成，对 $C[0,2\pi]$ 中的函数 f, W 中关于 f 的最佳逼近称为 f 的 **n 阶傅里叶逼近**. 由于（5）中的函数是正交的，给出的最佳逼近是 W 上的正交投影，在这种情形下，（4）式中的系数称为 f 的**傅里叶系数**. 标准的正交投影公式表明：

$$a_k=\frac{\langle f,\cos kt\rangle}{\langle \cos kt,\cos kt\rangle},\quad b_k=\frac{\langle f,\sin kt\rangle}{\langle \sin kt,\sin kt\rangle},\qquad k\geqslant 1$$

习题 7 要求你证明 $\langle \cos kt,\cos kt\rangle=\pi$ 和 $\langle \sin kt,\sin kt\rangle=\pi$，那么

$$a_k=\frac{1}{\pi}\int_0^{2\pi}f(t)\cos kt\mathrm{d}t,\quad b_k=\frac{1}{\pi}\int_0^{2\pi}f(t)\sin kt\mathrm{d}t \tag{7}$$

（常数）函数 1 的系数的正交投影是：

$$\frac{\langle f,1\rangle}{\langle 1,1\rangle}=\frac{1}{2\pi}\int_0^{2\pi}f(t)\cdot 1\mathrm{d}t=\frac{1}{2}\left[\frac{1}{\pi}\int_0^{2\pi}f(t)\cos(0\cdot t)\mathrm{d}t\right]=\frac{a_0}{2}$$

此处 a_0 是（7）式中 $k=0$ 的情形，这就解释了（4）中的常数项为什么写成 $a_0/2$.

例 4 求函数 $f(t)=t$ 在区间 $[0,2\pi]$ 上的 n 阶傅里叶逼近.

解 我们计算

$$\frac{a_0}{2}=\frac{1}{2}\cdot\frac{1}{\pi}\int_0^{2\pi}t\mathrm{d}t=\frac{1}{2\pi}\left[\frac{1}{2}t^2\Bigg|_0^{2\pi}\right]=\pi$$

当 $k>0$ 时，利用分部积分

$$a_k=\frac{1}{\pi}\int_0^{2\pi}t\cos kt\mathrm{d}t=\frac{1}{\pi}\left[\frac{1}{k^2}\cos kt+\frac{t}{k}\sin kt\right]_0^{2\pi}=0$$
$$b_k=\frac{1}{\pi}\int_0^{2\pi}t\sin kt\mathrm{d}t=\frac{1}{\pi}\left[\frac{1}{k^2}\sin kt-\frac{t}{k}\cos kt\right]_0^{2\pi}=-\frac{2}{k}$$

这样，$f(t)=t$ 的 n 阶傅里叶逼近是

$$\pi-2\sin t-\sin 2t-\frac{2}{3}\sin 3t-\cdots-\frac{2}{n}\sin nt$$

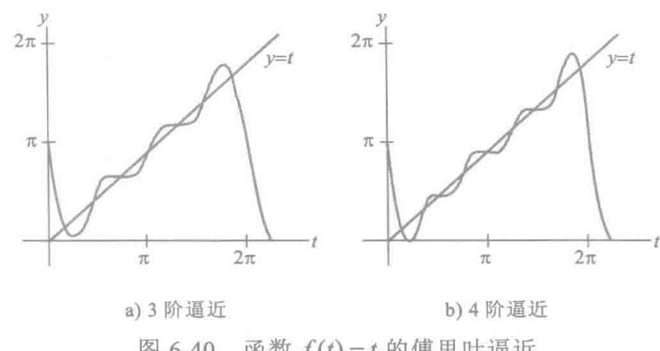
■

图 6-40 表示 f 的 3 阶和 4 阶傅里叶逼近.

函数 f 与傅里叶逼近之差的范数称为逼近的**均方误差**.（术语“均”是相对于积分定义的范数而言的.）可以证明，当傅里叶级数的阶数增加时，均方误差趋于零，由于这个原因，它常常写成

$$f(t)=\frac{a_0}{2}+\sum_{m=1}^{\infty}(a_m\cos mt+b_m\sin mt)$$

这个 $f(t)$ 的表达式称为 f 在 $[0,2\pi]$ 上的**傅里叶级数**. 例如，项 $a_m\cos mt$ 是 f 在由 $\cos mt$ 生成的一维子空间的投影.



a) 3 阶逼近　　b) 4 阶逼近

图 6-40　函数 $f(t)=t$ 的傅里叶逼近

练习题

1. 设 $q_1(t)=1$，$q_2(t)=t$ 和 $q_3(t)=3t^2-4$，验证 $\{q_1,q_2,q_3\}$ 是 $C[-2,2]$ 上的正交集，内积具有 6.7 节例 7 的形式（从 −2 到 2 积分）.
2. 求下列函数的 1 阶和 3 阶傅里叶逼近：

$$f(t)=3-2\sin t+5\sin 2t-6\cos 2t$$

习题 6.8

1. 求最小二乘直线 $y=\beta_0+\beta_1 x$，最佳拟合数据 $(-2,0),(-1,0),(0,2),(1,4)$ 和 $(2,4)$，若第一个和最后一个数据具有较小的可靠性，权值取中间三个数据的一半.
2. 假若加权最小二乘问题中，25 个数据里有 5 个数据具有的 y 度量且比其他数据的可靠性小，这些数据的权值被赋予其他 20 个数据权值的一半. 其中一个方法是 20 个数据的权值为 1，其余 5 个数据的权值为 $\frac{1}{2}$；另一个方法是 20 个数据权值为 2，其余 5 个数据的权值为 1. 这两种方法结果一致吗？试解释之.
3. 用三次趋势函数拟合例 2 中的数据，三次正交多项式是 $p_3(t)=\frac{5}{6}t^3-\frac{17}{6}t$.
4. 为给出 6 个等分数据点的趋势分析，可以用 $t=-5,-3,-1,1,3$ 和 5 的点观测值的正交多项式.
   a. 证明前三个正交多项式是

   $$p_0(t)=1,\ p_1(t)=t,\ p_2(t)=\frac{3}{8}t^2-\frac{35}{8}$$

   （多项式 p_2 已被重新度量，使得它在观测点处的值是小整数.）
   b. 利用二次趋势函数拟合数据 $(-5,1),(-3,1),(-1,4),(1,4),(3,6),(5,8)$.

在习题 5~14 中，空间 $C[0,2\pi]$ 具有（6）式的内积.

5. 证明：当 $m\neq n$ 时，$\sin mt$ 和 $\sin nt$ 正交.
6. 证明：对所有正整数 m 和 n, $\sin mt$ 和 $\cos nt$ 正交.
7. 证明：对 $k>0, \|\cos kt\|^2=\pi, \|\sin kt\|^2=\pi$.
8. 求函数 $f(t)=t-1$ 的 3 阶傅里叶逼近.
9. 求函数 $f(t)=2\pi-t$ 的 3 阶傅里叶逼近.
10. 求方波函数 $f(t)=\begin{cases}1 & 0\leqslant t<\pi\\ -1 & \pi\leqslant t<2\pi\end{cases}$ 的 3 阶傅里叶逼近.
11. 求 $\sin^2 t$ 的 3 阶傅里叶逼近，不要使用任何积分运算.
12. 求 $\cos^3 t$ 的 3 阶傅里叶逼近，不要使用任何积分运算.
13. 解释为什么两个函数和的傅里叶系数是两个函数傅里叶系数之和.

14. 假若 $C[0,2\pi]$ 中一些函数 f，前面几个傅里叶系数是 a_0,a_1,a_2 和 b_1,b_2,b_3．下面哪一个三角多项式更接近 f？解释你的答案.
$g(t)=\frac{a_0}{2}+a_1\cos t+a_2\cos 2t+b_1\sin t$
$h(t)=\frac{a_0}{2}+a_1\cos t+a_2\cos 2t+b_1\sin t+b_2\sin 2t$

15. [M]对 6.6 节习题 13 中的数据，涉及一个飞机的起飞表演. 假若随着飞机速度的增加，测量的误差可能变得更大，且 W 是加权对角矩阵且对角线的值是 1, 1 ,1, 0.9, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2 和 0.1，求一个立方曲线，使得拟合数据具有最小的加权最小二乘误差. 且利用结论估计飞机在 t=4.5 秒的速度.

16. [M] 对习题 10 中属于空间 $C[0,2\pi]$ 的方波函数，设 f_4 和 f_5 分别表示其 4 阶和 5 阶傅里叶逼近. 分别画出 f_4 和 f_5 在区间 $[0,2\pi]$ 上的图形和 f_5 在区间 $[-2\pi,2\pi]$ 上的图形.

练习题答案

1. 计算

$$\langle q_1,q_2\rangle=\int_{-2}^{2}1\cdot t\mathrm{d}t=\frac{1}{2}t^2\bigg|_{-2}^{2}=0$$

$$\langle q_1,q_3\rangle=\int_{-2}^{2}1\cdot(3t^2-4)\mathrm{d}t=(t^3-4t)\bigg|_{-2}^{2}=0$$

$$\langle q_2,q_3\rangle=\int_{-2}^{2}t\cdot(3t^2-4)\mathrm{d}t=\left(\frac{3}{4}t^4-2t^2\right)\bigg|_{-2}^{2}=0$$

2. f 的 3 阶傅里叶逼近是 $C[0,2\pi]$ 中 f 的最佳逼近，它是由 $1,\cos t,\cos 2t,\cos 3t,\sin t$, $\sin 2t$ 和 $\sin 3t$ 生成子空间中的向量函数. 显然 f 属于这个子空间，所以 f 是自己的最佳逼近.

$$f(t)=3-2\sin t+5\sin 2t-6\cos 2t$$

对 1 阶逼近，子空间 $W=\mathrm{Span}\{1,\cos t,\sin t\}$ 中最接近 f 的函数是 $3-2\sin t$， $f(t)$ 公式中的另外两项与 W 中的函数正交，所以，它们没有增加傅里叶系数中 1 阶逼近的积分计算.

见图 6-41.

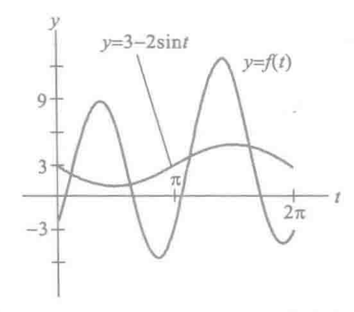


图 6-41 $f(t)$ 的 1 阶和 3 阶逼近

补充习题

1. 下列命题是针对属于 $\mathbb{R}^n$（或 $\mathbb{R}^m$）的具有标准内积的向量，确定每个命题的真假.

a. 每个向量的长度是一个正数.

b. 一个向量 $\boldsymbol{v}$ 与它的负向量 $-\boldsymbol{v}$ 具有同样长度.

c. $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 之间的距离是 $\|\boldsymbol{u}-\boldsymbol{v}\|$.

d. 如果 r 是任意数，那么 $\|r\boldsymbol{v}\|=r\|\boldsymbol{v}\|$.

e. 如果两个向量正交，则它们线性无关.

f. 如果 $\boldsymbol{x}$ 与 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 都正交，那么 $\boldsymbol{x}$ 一定与 $\boldsymbol{u}-\boldsymbol{v}$ 正交.

g. 如果 $\|\boldsymbol{u}+\boldsymbol{v}\|^2=\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2$，那么 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 正交.

h. 如果 $\|\boldsymbol{u}-\boldsymbol{v}\|^2=\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2$，那么 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 正交.

i. $\boldsymbol{y}$ 在 $\boldsymbol{u}$ 上的正交投影是一个数和 $\boldsymbol{y}$ 的乘积.

j. 如果向量 $\boldsymbol{y}$ 与它在 W 上的正交投影一致，那么 $\boldsymbol{y}$ 属于 W.

k. $\mathbb{R}^n$ 中所有正交于一个固定向量的集合是 $\mathbb{R}^n$ 的一个子空间.

l. 如果 W 是 $\mathbb{R}^n$ 的子空间，那么 W 和 $W^\perp$ 没有共同的向量.

m. 如果 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 是一个正交集，且如果 c_1,c_2,c_3 是数，那么 $\{c_1\boldsymbol{v}_1,c_2\boldsymbol{v}_2,c_3\boldsymbol{v}_3\}$ 也是一个正交集.

n. 如果一个矩阵 U 具有单位正交列，那么 $UU^{\mathrm{T}}=I$.

o. 一个具有正交列的方阵是一个正交矩阵.

p. 如果一个方阵具有单位正交列，那么它也有单位正交行.

q. 如果 W 是子空间，那么 $\|\mathrm{proj}_W\boldsymbol{v}\|^2+\|\boldsymbol{v}-\mathrm{proj}_W\boldsymbol{v}\|^2=\|\boldsymbol{v}\|^2$.

r. $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解是 $\mathrm{Col}\,A$ 中最接近 $\boldsymbol{b}$ 的向量 $A\hat{\boldsymbol{x}}$，即对所有的 $\boldsymbol{x}$ 有 $\|\boldsymbol{b}-A\hat{\boldsymbol{x}}\|\leqslant\|\boldsymbol{b}-A\boldsymbol{x}\|$.

s. $A\boldsymbol{x}=\boldsymbol{b}$ 最小二乘解对应法方程的解是 $\hat{\boldsymbol{x}}=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b}$.

2. 设 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 是一个单位正交集，用归纳法从 $p=2$ 开始验证，如果 $\boldsymbol{x}=c_1\boldsymbol{v}_1+\cdots+c_p\boldsymbol{v}_p$，那么 $\|\boldsymbol{x}\|^2=|c_1|^2+\cdots+|c_p|^2$.

3. 设 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 是 $\mathbb{R}^n$ 中单位正交集，验证下列贝塞尔不等式，对所有属于 $\mathbb{R}^n$ 的向量的正确性：
$$\|\boldsymbol{x}\|^2\geqslant|\boldsymbol{x}\cdot\boldsymbol{v}_1|^2+|\boldsymbol{x}\cdot\boldsymbol{v}_2|^2+\cdots+|\boldsymbol{x}\cdot\boldsymbol{v}_p|^2$$

4. 设 U 是 $n\times n$ 正交矩阵，证明：如果 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$ 是 $\mathbb{R}^n$ 中的一个单位正交基，那么 $\{U\boldsymbol{v}_1,\cdots,U\boldsymbol{v}_n\}$ 也是.

5. 证明：如果一个 $n\times n$ 矩阵 U，对所有属于 $\mathbb{R}^n$ 的向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$，满足 $(U\boldsymbol{x})\cdot(U\boldsymbol{y})=\boldsymbol{x}\cdot\boldsymbol{y}$，那么 U 是一个正交矩阵.

6. 证明：如果 U 是一个正交矩阵，那么 U 的任何实特征值一定是 ±1.

7. 一个豪斯霍尔德矩阵或基本镜像具有形式 $Q=I-2\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}$，此处 $\boldsymbol{u}$ 是一个单位向量（见第2章补充习题中的习题 13），证明 Q 是一个正交矩阵.（基本镜像经常用于在计算机程序中产生矩阵 A 的一个 QR 分解. 如果 A 具有线性无关的列，那么一系列基本镜像的左乘可产生一个上三角形矩阵.）

8. 设 $T:\mathbb{R}^n\to\mathbb{R}^n$ 为一保长的线性变换；即对 $\mathbb{R}^n$ 上所有 $\boldsymbol{x}$ 有 $\|T(\boldsymbol{x})\|=\|\boldsymbol{x}\|$.

 a. 证明：T 同时保持正交性；即只要 $\boldsymbol{x}\cdot\boldsymbol{y}=0$ 有 $T(\boldsymbol{x})\cdot T(\boldsymbol{y})=0$.

 b. 证明：T 的标准矩阵是一正交矩阵.

9. 设 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是 $\mathbb{R}^n$ 中线性无关的非正交的向量，描述如何不首先构造 $\mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$ 的正交基，可找出 $\mathbb{R}^n$ 中向量 $\boldsymbol{z}$ 的形如 $x_1\boldsymbol{u}+x_2\boldsymbol{v}$ 的最佳逼近.

10. 假若 A 的列是线性无关，确定当 $\boldsymbol{b}$ 用 $c\boldsymbol{b}$（c 是非零标量）代替时，$A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解 $\hat{\boldsymbol{x}}$ 会发生什么变化？

11. 如果 a,b,c 是不同的数，由于方程的图形是平行平面，使得下面的方程组不相容. 证明，该方程组的所有最小二乘解就是方程 $x-2y+5z=(a+b+c)/3$ 所确定的平面.
$$x-2y+5z=a$$
$$x-2y+5z=b$$
$$x-2y+5z=c$$

12. 已知逼近特征向量 $\boldsymbol{v}$，考虑如何求 $n\times n$ 矩阵 A 的特征值. 由于 $\boldsymbol{v}$ 不是准确向量，方程
$$A\boldsymbol{v}=\lambda\boldsymbol{v}\qquad(1)$$
可能没有解. 然而，可以通过适当观察（1）来求得最小二乘解来估计 λ. 把 $\boldsymbol{v}$ 看成 $n\times1$ 矩阵 V，把 λ 看作 $\mathbb{R}^1$ 上的一个向量，并用符号 $\boldsymbol{b}$ 表示向量 $A\boldsymbol{v}$. 从而（1）变成 $\boldsymbol{b}=\lambda V$，同样可写成 $V\lambda=\boldsymbol{b}$. 求由含有未知数 λ 的这 n 个方程组成的方程组的最小二乘解，并用原来符号表示出来. 对 λ 的估计结果叫做瑞利商. 参照 5.8 节中习题 11 和习题 12.

13. 利用下面步骤证明，下面 $m\times n$ 矩阵 A 的四个

基本子空间之间的联系.

$$\text{Row } A = (\text{Nul } A)^{\perp}, \text{Col } A = (\text{Nul } A^{\text{T}})^{\perp}$$

a. 证明 Row A 属于空间 $(\text{Nul } A)^{\perp}$.（证明如果 $\boldsymbol{x}$ 属 Row A，那么 $\boldsymbol{x}$ 与 Nul A 中任意一个 $\boldsymbol{u}$ 正交.）

b. 假若 rank $A = r$，找出 $\dim(\text{Nul } A)$ 和 $\dim(\text{Nul } A)^{\perp}$，然后从（a）推出 Row $A = (\text{Nul } A)^{\perp}$.（提示：研究 6.3 节的练习.）

c. 解释为什么 Col $A = (\text{Nul } A^{\text{T}})^{\perp}$.

14. 解释为什么方程 $A\boldsymbol{x} = \boldsymbol{b}$ 有解的充分必要条件是：$\boldsymbol{b}$ 与方程 $A^{\text{T}}\boldsymbol{x} = \boldsymbol{0}$ 的所有解正交.

习题 15 和习题 16 涉及一个 $n \times n$ 矩阵形如 $A = URU^{\text{T}}$ 的（实）舒尔分解，此处 U 是一个正交矩阵，R 是一个 $n \times n$ 上三角形矩阵.㊀

15. 证明：如果 A 有一个实舒尔分解，$A = URU^{\text{T}}$，那么 A 具有 n 个实特征值，计算包含重数.

16. 假若 A 是具有 n 个实特征根的 $n \times n$ 矩阵，包含重数，且记为 $\lambda_1, \cdots, \lambda_n$. 可以证明 A 有一个（实的） 舒尔分解，（a）和（b）两部分给出证明的关键思想，证明的其余部分等同于对小矩阵连续重复步骤（a）和（b），然后整合到一起得到最后结果.

a. 设 $\boldsymbol{u}_1$ 是对应于 λ_1 的单位特征向量，$\boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ 是其余向量且 $\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n\}$ 是 $\mathbb{R}^n$ 的单位正交基，则取 $U = [\boldsymbol{u}_1 \ \ \boldsymbol{u}_2 \ \cdots \ \boldsymbol{u}_n]$，证明 $U^{\text{T}}AU$ 的第一列是 $\lambda_1\boldsymbol{e}_1$，此处 $\boldsymbol{e}_1$ 是 $n \times n$ 单位矩阵的第一列.

b.（a）部分意味着 $U^{\text{T}}AU$ 具有下列的形式，解释为什么 A_1 的特征是 $\lambda_2, \cdots, \lambda_n$（提示：参考第 5 章的补充练习.）

$$U^{\text{T}}AU = \begin{bmatrix} \lambda_1 & * & * & * & * \\ 0 & & & & \\ \vdots & & A_1 & & \\ 0 & & & & \end{bmatrix}$$

[M]当方程 $A\boldsymbol{x} = \boldsymbol{b}$ 的右端有一点变化，例如 $A\boldsymbol{x} = \boldsymbol{b} + \Delta\boldsymbol{b}$，$\Delta\boldsymbol{b}$ 为向量，方程的解会由 $\boldsymbol{x}$ 改变为 $\boldsymbol{x} + \Delta\boldsymbol{x}$，此处 $\Delta\boldsymbol{x}$ 满足 $A(\Delta\boldsymbol{x}) = \Delta\boldsymbol{b}$. 商 $\|\Delta\boldsymbol{b}\| / \|\boldsymbol{b}\|$ 称为 **$\boldsymbol{b}$ 的相对改变**（或称 **$\boldsymbol{b}$ 的相对误差**，当 $\Delta\boldsymbol{b}$ 表示 $\boldsymbol{b}$ 的分量可能误差）. 解的相对改变是 $\|\Delta\boldsymbol{x}\| / \|\boldsymbol{x}\|$，当 A 可逆时，A 的**条件数**，记为 cond(A)，给出一个 $\boldsymbol{x}$ 相对改变多大的界：

$$\frac{\|\Delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leqslant \text{cond}(A) \cdot \frac{\|\Delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|} \qquad (2)$$

在习题 17~20 中，解出 $A\boldsymbol{x} = \boldsymbol{b}$ 和 $A(\Delta\boldsymbol{x}) = \Delta\boldsymbol{b}$，且证明（2）在每种情形下成立（参考 2.3 节习题 41~43 中关于病态矩阵的讨论）.

17. $A = \begin{bmatrix} 4.5 & 3.1 \\ 1.6 & 1.1 \end{bmatrix}$，$\boldsymbol{b} = \begin{bmatrix} 19.249 \\ 6.843 \end{bmatrix}$，$\Delta\boldsymbol{b} = \begin{bmatrix} 0.001 \\ -0.003 \end{bmatrix}$

18. $A = \begin{bmatrix} 4.5 & 3.1 \\ 1.6 & 1.1 \end{bmatrix}$，$\boldsymbol{b} = \begin{bmatrix} 0.500 \\ -1.407 \end{bmatrix}$，$\Delta\boldsymbol{b} = \begin{bmatrix} 0.001 \\ -0.003 \end{bmatrix}$

19. $A = \begin{bmatrix} 7 & -6 & -4 & 1 \\ -5 & 1 & 0 & -2 \\ 10 & 11 & 7 & -3 \\ 19 & 9 & 7 & 1 \end{bmatrix}$，$\boldsymbol{b} = \begin{bmatrix} 0.100 \\ 2.888 \\ -1.404 \\ 1.462 \end{bmatrix}$，

$\Delta\boldsymbol{b} = 10^{-4}\begin{bmatrix} 0.49 \\ -1.28 \\ 5.78 \\ 8.04 \end{bmatrix}$

20. $A = \begin{bmatrix} 7 & -6 & -4 & 1 \\ -5 & 1 & 0 & -2 \\ 10 & 11 & 7 & -3 \\ 19 & 9 & 7 & 1 \end{bmatrix}$，$\boldsymbol{b} = \begin{bmatrix} 4.230 \\ -11.043 \\ 49.991 \\ 69.536 \end{bmatrix}$，

$\Delta\boldsymbol{b} = 10^{-4}\begin{bmatrix} 0.27 \\ 7.76 \\ -3.77 \\ 3.93 \end{bmatrix}$

㊀ 如果允许复数，每一个 $n \times n$ 矩阵 A 有一个（复数）舒尔分解 $A = URU^{-1}$，此处 R 是上三角矩阵，U^{-1} 是 U 的共轭转置，这是 *Matrix Analysis*（Roger A. Horn and Charles R. Johnson, Cambridge: Cambridge University Press, 1985, pp. 79- 100）中一个非常重要的事实.

第7章　对称矩阵和二次型

介绍性实例　多波段的图像处理

在地球上，经过 80 分钟稍微多一点，两颗地球资源探测卫星静静地沿靠近极地的轨道飞越天空，以 185 公里宽的幅度，记录地形和海岸线的图像. 每颗卫星在每隔 16 天会扫遍地球的几乎每一平方公里，以致任何地方在 8 天内可被监测到.

地球资源探测卫星所拍到的图像有很多用途，研究人员和城市规划人员利用它们研究城市发展速度和方向，工业发展和土地使用的变化. 在乡村可用于分析土壤湿度，对偏远地区植被进行分类，确定内陆的湖泊和河流的位置. 政府部门可检测和评估自然灾害的破坏程度，如森林火灾、火山熔岩的流动、洪水和飓风等. 环保部门可以确定来自烟囱的污染，测量水力发电站附近的湖泊和河流的温度.

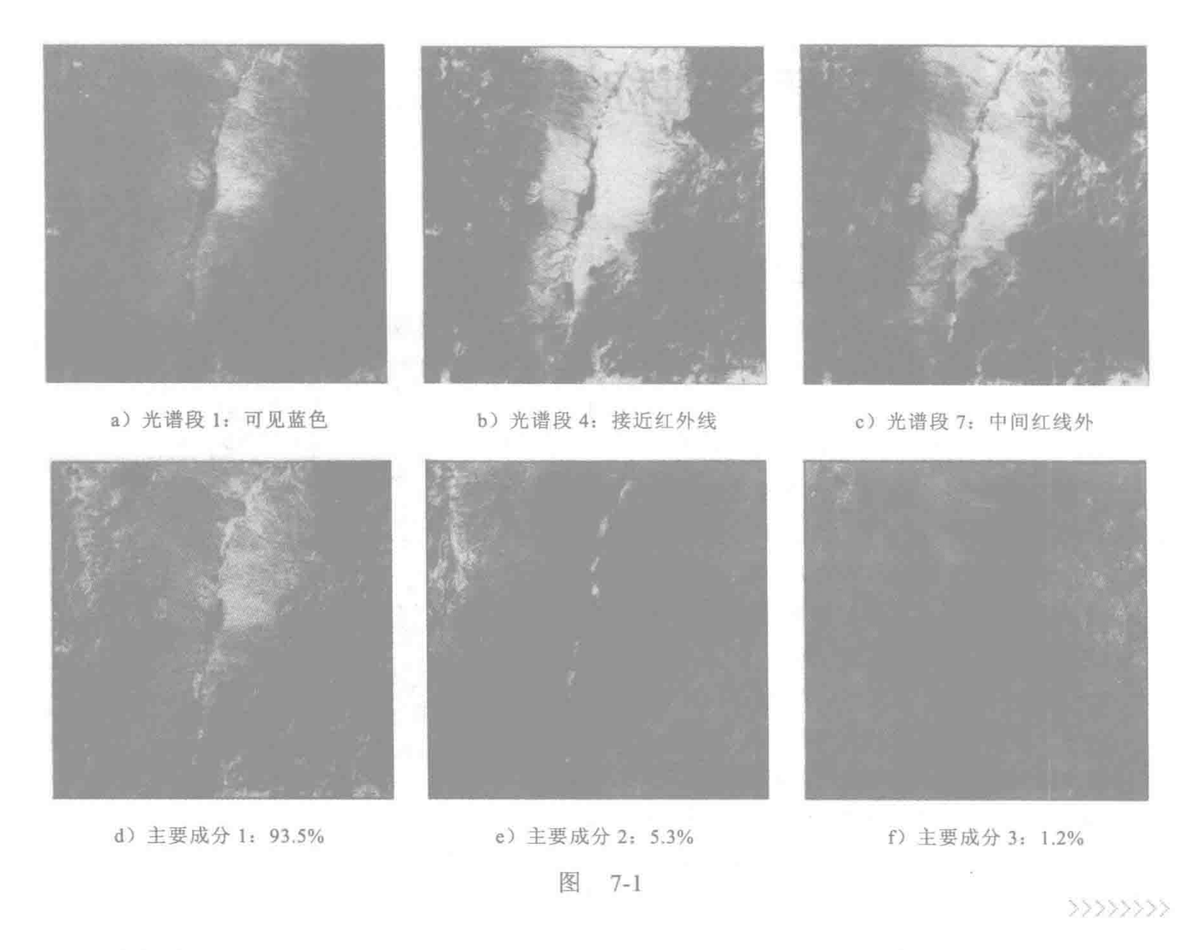
为了用于研究，卫星上的传感器可同步取得地球上任何地区的七种图像，传感器通过不同的波段来记录能量，包括三种可见光谱、四种红外光谱，每幅图像都被数字化且存储为矩阵，每一个数表示图像上对应点（像素）的信号强度，这七幅图像当中的每一幅都是多波段或多光谱图像的一个波段的影像.

一个固定区域的七幅地球资源探测卫星图像通常包含大量冗余的信息，原因是一些特性会表现在几幅图像中. 然而其他特性，由于它们的颜色或温度，会反射出只被一个或两个传感器记录的光线. 多波段图像处理的一个目标是，用一种比研究每幅图像更好的方式来提取信息去观察数据.

主成分分析是一种从原始数据中消除冗余信息，因而只需要一幅或两幅合成图像就可以提供大部分信息的有效方法. 粗略地说，其目的是找出一个特殊的图像线性组合，即给七种像素中每一个赋予权值，然后再综合得到一个新的图像值. 权值选取的方式使得合成图像中的光线强度的变化幅度或*景象差异*（称为*第一主成分*）比任何原始图像的都要大. 更多成分图像也可用一定的准则来构造，7.5 节会给出准则的解释.

主成分分析也可用取自内华达州铁路峡谷的图像（见图 7-1）来解释，地球资源探测卫星三个波段拍到的图像是图 a~c. 三个波段的所有信息被重新组合成三个主成分图像，如图 d~f. 第一个成分 d 显示（或解释）93.5%的原始数据中的景象差异. 用这种方式，三个波段的原始数据被重新组合成一个波段的数据，且在某种意义下的景象差异仅损失 6.5%.

马里兰州的罗克维尔地球卫星公司，友好地提供了显示在这里的图片，他们正在对 224 种不同波段记录的图像进行合成图像的实验. 面对如此大量的数据，使用主成分分析，一般可以将数据减少到约 15 个有用的主成分.



a）光谱段 1：可见蓝色　b）光谱段 4：接近红外线　c）光谱段 7：中间红线外

d）主要成分 1：93.5%　e）主要成分 2：5.3%　f）主要成分 3：1.2%

图 7-1

与其他类型的矩阵相比，对称矩阵比其他类型的矩阵更常出现在应用中，其理论完美，本质上依赖于第 5 章的对角化技巧和第 6 章的正交性. 7.1 节叙述的对称矩阵对角化，是 7.2 节和 7.3 节讨论二次型的基础，最后两节讨论的奇异值分解和实例介绍中所描述的图像处理，则需要 7.3 节的内容.在整章中，所有向量和矩阵的元素均为实数.

7.1 对称矩阵的对角化

一个对称矩阵是一个满足 $A^{\mathrm{T}}=A$ 的矩阵 A，这种矩阵当然是方阵，它的主对角线元素是任意的，但其他元素在主对角线的两边成对出现.

例 1 下面的矩阵中，仅前面三个是对称矩阵.

$$\text{对称：}\begin{bmatrix}1 & 0\\ 0 & -3\end{bmatrix},\quad \begin{bmatrix}0 & -1 & 0\\ -1 & 5 & 8\\ 0 & 8 & -7\end{bmatrix},\quad \begin{bmatrix}a & b & c\\ b & d & e\\ c & e & f\end{bmatrix}$$

$$\text{非对称：}\begin{bmatrix}1 & -3\\3 & 0\end{bmatrix},\quad \begin{bmatrix}1 & -4 & 0\\-6 & 1 & -4\\0 & -6 & 1\end{bmatrix},\quad \begin{bmatrix}5 & 4 & 3 & 2\\4 & 3 & 2 & 1\\3 & 2 & 1 & 0\end{bmatrix}$$

■

为了开始学习对称矩阵，复习 5.3 节的对角化过程非常有用.

例 2 如果可能，对角化矩阵 $A=\begin{bmatrix}6 & -2 & -1\\-2 & 6 & -1\\-1 & -1 & 5\end{bmatrix}$.

解 A 的特征方程是

$$0=-\lambda^3+17\lambda^2-90\lambda+144=-(\lambda-8)(\lambda-6)(\lambda-3)$$

通过标准计算可得到每个特征子空间的一个基：

$$\lambda=8:\boldsymbol{v}_1=\begin{bmatrix}-1\\1\\0\end{bmatrix};\quad \lambda=6:\boldsymbol{v}_2=\begin{bmatrix}-1\\-1\\2\end{bmatrix};\quad \lambda=3:\boldsymbol{v}_3=\begin{bmatrix}1\\1\\1\end{bmatrix}$$

这三个向量形成 $\mathbb{R}^3$ 的一个基，事实上，很容易验证 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 是 $\mathbb{R}^3$ 的正交基. 根据第 6 章的经验，正交基对计算可能有用，因此可以单位化 $\boldsymbol{v}_1,\boldsymbol{v}_2$和$\boldsymbol{v}_3$ 得到单位特征向量.

$$\boldsymbol{u}_1=\begin{bmatrix}-1/\sqrt{2}\\1/\sqrt{2}\\0\end{bmatrix},\quad \boldsymbol{u}_2=\begin{bmatrix}-1/\sqrt{6}\\-1/\sqrt{6}\\2/\sqrt{6}\end{bmatrix},\quad \boldsymbol{u}_3=\begin{bmatrix}1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{bmatrix}$$

令

$$P=\begin{bmatrix}-1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3}\\1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3}\\0 & 2/\sqrt{6} & 1/\sqrt{3}\end{bmatrix},\quad D=\begin{bmatrix}8 & 0 & 0\\0 & 6 & 0\\0 & 0 & 3\end{bmatrix}$$

那么有 $A=PDP^{-1}$，和平常一样. 由于 P 是方阵且有正交列，所以，P 是一个正交矩阵，而 P^{-1} 就是 P^{T}.（见 6.2 节.） ■

定理 1 解释了为什么例 2 中的特征向量是正交的——它们对应不同的特征值.

定理 1 如果 A 是对称矩阵，那么不同特征空间的任意两个特征向量是正交的.

证 设 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 是对应不同特征值 λ_1,λ_2 的特征向量. 为证明 $\boldsymbol{v}_1\cdot\boldsymbol{v}_2=0$，计算

$$\begin{aligned}\lambda_1\boldsymbol{v}_1\cdot\boldsymbol{v}_2&=(\lambda_1\boldsymbol{v}_1)^{\mathrm{T}}\boldsymbol{v}_2=(A\boldsymbol{v}_1)^{\mathrm{T}}\cdot\boldsymbol{v}_2 &&\text{由于 } \boldsymbol{v}_1 \text{ 是一个特征向量}\\&=(\boldsymbol{v}_1^{\mathrm{T}}A^{\mathrm{T}})\boldsymbol{v}_2=\boldsymbol{v}_1^{\mathrm{T}}(A\boldsymbol{v}_2) &&\text{由于 } A^{\mathrm{T}}=A\\&=\boldsymbol{v}_1^{\mathrm{T}}(\lambda_2\boldsymbol{v}_2) &&\text{由于 } \boldsymbol{v}_2 \text{ 是一个特征向量}\\&=\lambda_2\boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{v}_2=\lambda_2\boldsymbol{v}_1\cdot\boldsymbol{v}_2\end{aligned}$$

因此 $(\lambda_1-\lambda_2)\boldsymbol{v}_1\cdot\boldsymbol{v}_2=0$，但是$\lambda_1\neq\lambda_2$，所以$\boldsymbol{v}_1\cdot\boldsymbol{v}_2=0$. ■

例 2 中特殊类型的对角化对于对称矩阵理论十分重要. 一个矩阵 A 称为可正交对角化，如

果存在一个正交矩阵 P（满足 $P^{-1}=P^{\mathrm{T}}$）和一个对角矩阵 D 使得：

$$A=PDP^{\mathrm{T}}=PDP^{-1} \tag{1}$$

为了正交对角化一个 $n\times n$ 矩阵，我们必须找到 n 个线性无关且单位正交的特征向量，什么条件下可能做到？如果 A 是像（1）一样可以正交对角化的，那么

$$A^{\mathrm{T}}=(PDP^{\mathrm{T}})^{\mathrm{T}}=P^{\mathrm{TT}}D^{\mathrm{T}}P^{\mathrm{T}}=PDP^{\mathrm{T}}=A$$

这样 A 是对称的！另一方面，定理 2 表明每一个对称矩阵都是可正交对角化，结论的证明非常困难，这里将其省略，证明的主要思路将在定理 3 之后给出.

定理 2 *一个 $n\times n$ 矩阵 A 可正交对角化的充分必要条件是 A 是对称矩阵.*

这个定理相当奇妙，我们根据第 5 章的经验可以推测，要知道何时一个矩阵 A 是可对角化通常是不可能的，但对称矩阵却例外.

下面的例子给出一个矩阵，它的特征值并非全部都不相同.

例 3 将 $A=\begin{bmatrix}3&-2&4\\-2&6&2\\4&2&3\end{bmatrix}$ 正交对角化，其特征方程为

$$0=-\lambda^3+12\lambda^2-21\lambda-98=-(\lambda-7)^2(\lambda+2)$$

解 平常计算可得特征空间的基：

$$\lambda=7:\boldsymbol{v}_1=\begin{bmatrix}1\\0\\1\end{bmatrix},\ \boldsymbol{v}_2=\begin{bmatrix}-1/2\\1\\0\end{bmatrix},\ \lambda=-2:\boldsymbol{v}_3=\begin{bmatrix}-1\\-1/2\\1\end{bmatrix}$$

尽管 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 是线性无关的，但它们并不正交. 从 6.2 节可知 $\boldsymbol{v}_2$ 在 $\boldsymbol{v}_1$ 上的正交投影是 $\dfrac{\boldsymbol{v}_2\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1}\cdot\boldsymbol{v}_1$，与 $\boldsymbol{v}_1$ 正交的关于 $\boldsymbol{v}_2$ 的分量是：

$$\boldsymbol{z}_2=\boldsymbol{v}_2-\frac{\boldsymbol{v}_2\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1}\boldsymbol{v}_1=\begin{bmatrix}-1/2\\1\\0\end{bmatrix}-\frac{-1/2}{2}\begin{bmatrix}1\\0\\1\end{bmatrix}=\begin{bmatrix}-1/4\\1\\1/4\end{bmatrix}$$

于是 $\{\boldsymbol{v}_1,\boldsymbol{z}_2\}$ 是关于 $\lambda=7$ 的特征空间的正交集.（注意 $\boldsymbol{z}_2$ 是 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的线性组合，则 $\boldsymbol{z}_2$ 属于特征子空间，构造 $\boldsymbol{z}_2$ 的过程就是 6.4 节的格拉姆-施密特方法.）由于特征子空间是二维的（基是 $\boldsymbol{v}_1,\boldsymbol{v}_2$），正交集合 $\{\boldsymbol{v}_1,\boldsymbol{z}_2\}$ 是一个特征子空间的正交基（见 2.9 节或 4.5 节）.

将 $\boldsymbol{v}_1$ 和 $\boldsymbol{z}_2$ 规范化，我们得到关于 $\lambda=7$ 特征子空间的单位正交基.

$$\boldsymbol{u}_1=\begin{bmatrix}1/\sqrt{2}\\0\\1/\sqrt{2}\end{bmatrix},\ \boldsymbol{u}_2=\begin{bmatrix}-1/\sqrt{18}\\4/\sqrt{18}\\1/\sqrt{18}\end{bmatrix}$$

关于 $\lambda=-2$ 对应的特征空间的单位正交基是

$$u_3 = \frac{1}{\|2v_3\|} \cdot 2v_3 = \frac{1}{3}\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$$

由定理 1，u_3 与其他特征向量 u_1 和 u_2 正交，因此 $\{u_1, u_2, u_3\}$ 是一个正交集. 令

$$P = [u_1\ u_2\ u_3] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \\ 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \end{bmatrix}, D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

那么 P 将 A 正交对角化且 $A = PDP^{-1}$. ■

在例 3 中，特征值 7 是二重的，特征向量是二维的. 这个事实并不是偶然的，见下面定理.

谱定理

矩阵 A 的特征值的集合有时称为 A 的谱，且下面关于 A 的特征值描述称为谱定理.

定理 3 （对称矩阵的谱定理）

一个对称的 $n\times n$ 矩阵具有下面特性：

a. A 有 n 个实特征值，包含重复的特征值.

b. 对每一个特征值，对应特征子空间的维数等于 λ 作为特征方程的根的重数.

c. 特征空间相互正交，这种正交性是在特征向量对应不同特征值的意义下成立的.

d. A 可正交对角化.

（a）可从 5.5 节的习题 24 得出，（b）可从（d）容易得到（见练习 31）.（c）就是定理 1. 由于（a），（d）的一个证明可用习题 32 和第 6 章补充习题 16 已经讨论过的舒尔因子分解给出，证明的细节从略.

谱分解

假设 $A = PDP^{-1}$，此处 P 的列是 A 的单位正交特征向量 $u_1,\cdots,u_n$，且相应的特征值 $\lambda_1,\cdots,\lambda_n$ 属于对角矩阵 D，那么 $P^{-1} = P^{\mathrm{T}}$，

$$A = PDP^{\mathrm{T}} = [u_1 \cdots u_n]\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}\begin{bmatrix} u_1^{\mathrm{T}} \\ \vdots \\ u_n^{\mathrm{T}} \end{bmatrix}$$

$$= [\lambda_1 u_1 \cdots \lambda_n u_n]\begin{bmatrix} u_1^{\mathrm{T}} \\ \vdots \\ u_n^{\mathrm{T}} \end{bmatrix}$$

利用乘积的行列展开（2.4 节定理 10），我们可以得到

$$A = \lambda_1 u_1 u_1^{\mathrm{T}} + \lambda_2 u_2 u_2^{\mathrm{T}} + \cdots + \lambda_n u_n u_n^{\mathrm{T}} \tag{2}$$

由于它将 A 分解为 A 的谱（特征值）确定的小块，这个 A 的表示就称为 A 的**谱分解**，（2）中的每一项都是一个秩为 1 的 $n\times n$ 矩阵. 例如，$\lambda u_1 u_1^{\mathrm{T}}$ 的每一列都是 u_1 的倍数. 更进一步，在 x 属于 $\mathbb{R}^n$ 的意义下，矩阵 $u_j u_j^{\mathrm{T}}$ 是投影矩阵，向量 $(u_j u_j^{\mathrm{T}})x$ 是 x 在由 u_j 生成子空间上的正交投影.（见习题 35.）

例 4 构造矩阵 A 的一个谱分解，已知 A 有以下正交对角化分解：

$$A=\begin{bmatrix}7&2\\2&4\end{bmatrix}=\begin{bmatrix}2/\sqrt{5}&-1/\sqrt{5}\\1/\sqrt{5}&2/\sqrt{5}\end{bmatrix}\begin{bmatrix}8&0\\0&3\end{bmatrix}\begin{bmatrix}2/\sqrt{5}&1/\sqrt{5}\\-1/\sqrt{5}&2/\sqrt{5}\end{bmatrix}$$

解 将 P 的列记为 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$，则有

$$A=8\boldsymbol{u}_1\boldsymbol{u}_1^{\mathrm{T}}+3\boldsymbol{u}_2\boldsymbol{u}_2^{\mathrm{T}}$$

为验证 A 的谱分解，计算

$$\boldsymbol{u}_1\boldsymbol{u}_1^{\mathrm{T}}=\begin{bmatrix}2/\sqrt{5}\\1/\sqrt{5}\end{bmatrix}\begin{bmatrix}2/\sqrt{5}&1/\sqrt{5}\end{bmatrix}=\begin{bmatrix}4/5&2/5\\2/5&1/5\end{bmatrix}$$

$$\boldsymbol{u}_2\boldsymbol{u}_2^{\mathrm{T}}=\begin{bmatrix}-1/\sqrt{5}\\2/\sqrt{5}\end{bmatrix}\begin{bmatrix}-1/\sqrt{5}&2/\sqrt{5}\end{bmatrix}=\begin{bmatrix}1/5&-2/5\\-2/5&4/5\end{bmatrix}$$

且

$$8\boldsymbol{u}_1\boldsymbol{u}_1^{\mathrm{T}}+3\boldsymbol{u}_2\boldsymbol{u}_2^{\mathrm{T}}=\begin{bmatrix}32/5&16/5\\16/5&8/5\end{bmatrix}+\begin{bmatrix}3/5&-6/5\\-6/5&12/5\end{bmatrix}=\begin{bmatrix}7&2\\2&4\end{bmatrix}=A$$

■

数值计算的注解 当 A 是对称且不太大的矩阵时，现代高性能的计算机程序可以非常精确地计算特征值和特征向量，它们对 A 使用一系列包含正交矩阵的相似变换，变换后矩阵的对角元素很快收敛于 A 的特征值.（见 5.2 节的数值计算注释.）利用正交矩阵常常可避免计算过程的误差积累，当 A 对称时，正交矩阵序列可形成列向量是 A 的特征向量的正交矩阵.

一个非对称矩阵没有完全的正交特征向量集，但通过算法仍得到相当精确的特征值，之后，就需要用非正交化方法计算特征向量.

练习题

1. 证明：如果 A 是对称矩阵，那么 A^2 也是对称的.
2. 证明：如果 A 是可正交对角化的，那么 A^2 也是可正交对角化的.

习题 7.1

判断习题 1~6 中哪一个矩阵是对称矩阵.

1. $\begin{bmatrix}3&5\\5&-7\end{bmatrix}$

2. $\begin{bmatrix}-3&5\\-5&3\end{bmatrix}$

3. $\begin{bmatrix}2&2\\4&4\end{bmatrix}$

4. $\begin{bmatrix}0&8&3\\8&0&-2\\3&-2&0\end{bmatrix}$

5. $\begin{bmatrix}-6&2&0\\0&-6&2\\0&0&-6\end{bmatrix}$

6. $\begin{bmatrix}1&2&1&2\\2&1&2&1\\1&2&1&2\end{bmatrix}$

判断习题 7~12 中哪一个矩阵是正交的，如果正交，求出它的逆矩阵.

7. $\begin{bmatrix}0.6&0.8\\0.8&-0.6\end{bmatrix}$

8. $\begin{bmatrix}1/\sqrt{2}&-1/\sqrt{2}\\1/\sqrt{2}&1/\sqrt{2}\end{bmatrix}$

9. $\begin{bmatrix}-5&2\\2&5\end{bmatrix}$

10. $\begin{bmatrix}-1&2&2\\2&-1&2\\2&2&-1\end{bmatrix}$

11. $\begin{bmatrix}2/3&2/3&1/3\\0&1/\sqrt{5}&-2/\sqrt{5}\\\sqrt{5}/3&-4/\sqrt{45}&-2/\sqrt{45}\end{bmatrix}$

12. $\begin{bmatrix} 0.5 & 0.5 & -0.5 & -0.5 \\ -0.5 & 0.5 & -0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 & 0.5 \\ -0.5 & 0.5 & 0.5 & -0.5 \end{bmatrix}$

将习题 13~22 中的矩阵正交对角化，并给出一个正交矩阵 P 和一个对角矩阵 D，为节约时间，给出习题 17~24 的特征值是：（17）5, 2，−2；（18）25，3，−50，（19）7，−2；（20）13，7，1，（21）9，5，1；（22）2，0.

13. $\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$

14. $\begin{bmatrix} 1 & 5 \\ 5 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 16 & -4 \\ -4 & 1 \end{bmatrix}$

16. $\begin{bmatrix} -7 & 24 \\ 24 & 7 \end{bmatrix}$

17. $\begin{bmatrix} 1 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 1 \end{bmatrix}$

18. $\begin{bmatrix} -2 & -36 & 0 \\ -36 & -23 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

19. $\begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$

20. $\begin{bmatrix} 7 & -4 & 4 \\ -4 & 5 & 0 \\ 4 & 0 & 9 \end{bmatrix}$

21. $\begin{bmatrix} 4 & 1 & 3 & 1 \\ 1 & 4 & 1 & 3 \\ 3 & 1 & 4 & 1 \\ 1 & 3 & 1 & 4 \end{bmatrix}$

22. $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$

23. 设 $A=\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$ 和 $v=\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$. 证明 2 是 A 的一个特征值且 v 是一个特征向量，然后将 A 正交对角化.

24. 设 $A=\begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & 2 \\ -2 & 2 & 2 \end{bmatrix}$，$v_1=\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$ 和 $v_2=\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$. 验证 v_1 和 v_2 是 A 的特征向量，然后将 A 正交对角化.

对习题 25~26，判断每一个命题的真假，验证你的答案.

25. a. 可正交对角化的矩阵 A 一定是对称的.
 b. 如果 $A^{\mathrm{T}}=A$ 且向量 u 和 v 满足 $Au=3u$ 和 $Av=4v$，那么 $u\cdot v=0$.
 c. 一个 $n\times n$ 对称矩阵 A 有 n 个不同实特征值.
 d. 对属于 $\mathbb{R}^n$ 的非零向量 v，矩阵 vv^{T} 被称为投影矩阵.

26. a. 每一个对称矩阵 A 都是可正交对角化的.
 b. 如果 $B=PDP^{\mathrm{T}}$，此处 $P^{\mathrm{T}}=P^{-1}$ 且 D 是对角矩阵，那么 B 是对称矩阵.
 c. 一个正交矩阵是可正交对角化的.
 d. 一个对称矩阵的特征空间的维数等于对应特征值的重数.

27. 若 A 是对称的 $n\times n$ 矩阵，B 是任意 $n\times m$ 矩阵，证明 $B^{\mathrm{T}}AB$，$B^{\mathrm{T}}B$ 和 BB^{T} 是对称矩阵.

28. 证明：如果 A 是一个 $n\times n$ 对称矩阵，那么对任意属于 $\mathbb{R}^n$ 的 x, y 有 $(Ax)\cdot y = x\cdot(Ay)$ 成立.

29. 若 A 可逆且可正交对角化，解释为什么 A^{-1} 也是可正交对角化的.

30. 若 A 和 B 都可正交对角化且 $AB=BA$，解释为什么 AB 也是可正交对角化的.

31. 若 $A=PDP^{-1}$，其中 P 是正交矩阵，D 是对角矩阵，且 λ 是 A 的重数为 k 的特征值，那么 λ 在对角矩阵 D 中出现 k 次，解释为什么 λ 的特征空间的维数是 k.

32. 若 $A=PRP^{-1}$，其中 P 是正交的且 R 是上三角形矩阵，证明：如果 A 对称，那么 R 是对称的，因而实际上是一个对角矩阵.

33. 构造例 2 中矩阵 A 的一个谱分解.

34. 构造例 3 中矩阵 A 的一个谱分解.

35. 若 u 是 $\mathbb{R}^n$ 中的单位向量，取 $B=uu^{\mathrm{T}}$.
 a. 对任意属于 $\mathbb{R}^n$ 的 x，计算 Bx 且证明 Bx 是 x 在 u 上的正交投影，正如 6.2 节所描述的情况.
 b. 证明 B 是对称矩阵且 $B^2=B$.
 c. 证明 u 是 B 的特征向量，并求其对应的特征值.

36. 设 B 是 $n\times n$ 对称矩阵且 $B^2=B$，任何此类矩阵被称为**投影矩阵**（或一个**正交投影矩阵**）. 对任意给定的 $y\in\mathbb{R}^n$，取 $\hat{y}=By$ 且 $z=y-\hat{y}$.
 a. 证明 z 与 $\hat{y}$ 正交.

b．设 W 是 B 的列空间，证明 $\boldsymbol{y}$ 是 W 空间中一个向量与空间 $W^{\perp}$ 中一个向量之和. 证明 $B\boldsymbol{y}$ 是 $\boldsymbol{y}$ 在 B 的列空间上的正交投影.

[M]正交对角化习题 37~40 中的矩阵，为练习本节的方法，不要用矩阵程序中特征向量的常规解法，而是使用程序求出特征值，然后对每个特征值求出 $\text{Nul}(A-\lambda I)$ 的单位正交基，即类似例 2 和例 3 中的做法.

37. $\begin{bmatrix} 5 & 2 & 9 & -6 \\ 2 & 5 & -6 & 9 \\ 9 & -6 & 5 & 2 \\ -6 & 9 & 2 & 5 \end{bmatrix}$

38. $\begin{bmatrix} 0.38 & -0.18 & -0.06 & -0.04 \\ -0.18 & 0.59 & -0.04 & 0.12 \\ -0.06 & -0.04 & 0.47 & -0.12 \\ -0.04 & 0.12 & -0.12 & 0.41 \end{bmatrix}$

39. $\begin{bmatrix} 0.31 & 0.58 & 0.08 & 0.44 \\ 0.58 & -0.56 & 0.44 & -0.58 \\ 0.08 & 0.44 & 0.19 & -0.08 \\ 0.44 & -0.58 & -0.08 & 0.31 \end{bmatrix}$

40. $\begin{bmatrix} 10 & 2 & 2 & -6 & 9 \\ 2 & 10 & 2 & -6 & 9 \\ 2 & 2 & 10 & -6 & 9 \\ -6 & -6 & -6 & 26 & 9 \\ 9 & 9 & 9 & 9 & -19 \end{bmatrix}$

练习题答案

1. 由转置矩阵的性质得到 $(A^2)^{\mathrm{T}}=(AA)^{\mathrm{T}}=A^{\mathrm{T}}A^{\mathrm{T}}$，由假设，$A^{\mathrm{T}}=A$，所以 $(A^2)^{\mathrm{T}}=AA=A^2$，从而证明了 A^2 是对称的.
2. 如果 A 可正交对角化，那么由定理 2 可知矩阵 A 对称，由练习题 1，A^2 对称且可正交对角化（定理 2）.

7.2 二次型

到目前为止，本教材除了第 6 章计算 $x^{\mathrm{T}}x$ 时所遇到的平方和外，我们所关注的主要是线性方程，这类平方和及更一般形式的表达式称为二次型，它常常出现在线性代数在工程（标准设计和优化）和信号处理（输出的噪声功率）的应用中. 它们也常常出现在物理学（例如势能和动能）、微分几何（例如曲面的法曲率）、经济学（例如效用函数）和统计学（例如置信椭圆体）中，某些这类应用实例的数学背景很容易转化为对对称矩阵的研究.

$\mathbb{R}^n$ 上的一个**二次型**是一个定义在 $\mathbb{R}^n$ 上的函数，它在向量 $\boldsymbol{x}$ 处的值可由表达式 $Q(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 计算，此处 A 是一个 $n\times n$ 对称矩阵，且矩阵 A 称为**关于二次型的矩阵**.

最简单的非零二次型是 $Q(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}I\boldsymbol{x}=\|\boldsymbol{x}\|^2$，例 1 和例 2 说明了对称矩阵 A 和二次型 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 之间的关系.

例 1 令 $\boldsymbol{x}=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$，计算下列矩阵的 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$.

a．$A=\begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$　　　b．$A=\begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$

解

a．$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=[x_1 \quad x_2]\begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=[x_1 \quad x_2]\begin{bmatrix} 4x_1 \\ 3x_2 \end{bmatrix}=4x_1^{\ 2}+3x_2^{\ 2}$.

b．在 A 中有两个值为–2 的元素，注意观察矩阵 A 中（1,2）元素如何出现在计算中，用粗体字给出．

$$\begin{aligned}\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} &= [x_1 \quad x_2]\begin{bmatrix}3 & \mathbf{-2} \\ -2 & 7\end{bmatrix}\begin{bmatrix}x_1 \\ x_2\end{bmatrix} = [x_1 \quad x_2]\begin{bmatrix}3x_1 - \mathbf{2}x_2 \\ -2x_1 + 7x_2\end{bmatrix} \\ &= x_1(3x_1 - \mathbf{2}x_2) + x_2(-2x_1 + 7x_2) \\ &= 3x_1^2 - \mathbf{2}x_1x_2 - 2x_2x_1 + 7x_2^2 \\ &= 3x_1^2 - 4x_1x_2 + 7x_2^2\end{aligned}$$

■

例 1（b）中二次型中出现了项 $-4x_1x_2$，是因为矩阵 A 对角线上出现了元素–2，相反，例 1（a）中涉及的对角矩阵的二次型中没有 x_1x_2 的交叉乘积项．

例 2 对属于 $\mathbb{R}^3$ 的 $\boldsymbol{x}$，取 $Q(\boldsymbol{x}) = 5x_1^2 + 3x_2^2 + 2x_3^2 - x_1x_2 + 8x_2x_3$，写出 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 形式的二次型．

解 x_1^2, x_2^2, x_3^2 的系数仍在对角线上，为使 A 对称，当 $i \neq j$ 时，x_ix_j 的系数必须平均分配给矩阵 A 中的 (i,j) 元素和 (j,i) 元素．x_1x_3 的系数 0，很容易验证

$$Q(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} = [x_1 \quad x_2 \quad x_3]\begin{bmatrix}5 & -1/2 & 0 \\ -1/2 & 3 & 4 \\ 0 & 4 & 2\end{bmatrix}\begin{bmatrix}x_1 \\ x_2 \\ x_3\end{bmatrix}$$

■

例 3 令 $Q(\boldsymbol{x}) = x_1^2 - 8x_1x_2 - 5x_2^2$，计算 $Q(\boldsymbol{x})$ 在 $\boldsymbol{x} = \begin{bmatrix}-3 \\ 1\end{bmatrix}, \begin{bmatrix}2 \\ -2\end{bmatrix}$ 和 $\begin{bmatrix}1 \\ -3\end{bmatrix}$ 处的值．

解

$$\begin{aligned}Q(-3,\ 1) &= (-3)^2 - 8(-3)(1) - 5(1)^2 = 28 \\ Q(2,-2) &= (2)^2 - 8(2)(-2) - 5(-2)^2 = 16 \\ Q(1,-3) &= (1)^2 - 8(1)(-3) - 5(-3)^2 = -20\end{aligned}$$

■

在某些情况下，没有交叉项的二次型会显得更容易使用，也就是说，二次型对应的矩阵是对角矩阵．幸运的是，交叉项可以通过用适当的变量代换来消去．

二次型的变量代换

如果 $\boldsymbol{x}$ 表示 $\mathbb{R}^n$ 中的向量变量，那么**变量代换**是下面形式的等式：

$$\boldsymbol{x} = P\boldsymbol{y} \quad 或 \quad \boldsymbol{y} = P^{-1}\boldsymbol{x} \tag{1}$$

此处 P 是可逆矩阵且 $\boldsymbol{y}$ 是 $\mathbb{R}^n$ 中的一个新变量，这里 P 的列可确定 $\mathbb{R}^n$ 的一个基，$\boldsymbol{y}$ 是相对于该基向量 $\boldsymbol{x}$ 的坐标向量．（见 4.4 节．）

如果用变量代换（1）处理二次型，那么

$$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} = (P\boldsymbol{y})^{\mathrm{T}}A(P\boldsymbol{y}) = \boldsymbol{y}^{\mathrm{T}}P^{\mathrm{T}}AP\boldsymbol{y} = \boldsymbol{y}^{\mathrm{T}}(P^{\mathrm{T}}AP)\boldsymbol{y} \tag{2}$$

且新的二次型矩阵是 $P^{\mathrm{T}}AP$．因为 A 是对称的，由定理 2，存在正交矩阵 P，使得 $P^{\mathrm{T}}AP$ 是对角矩阵 D，（2）中的二次型矩阵变为 $\boldsymbol{y}^{\mathrm{T}}D\boldsymbol{y}$，这就是下面例题的解题思路．

例 4 求一个变量代换将例 3 中的二次型变为一个没有交叉项的二次型．

解 例 3 中二次型对应的矩阵是

$$A=\begin{bmatrix}1 & -4\\ -4 & -5\end{bmatrix}$$

第一步是将矩阵 A 正交对角化，A 的特征值是 $\lambda=3$ 和 $\lambda=-7$，相应的单位特征向量是：

$$\lambda=3:\begin{bmatrix}2/\sqrt{5}\\ -1/\sqrt{5}\end{bmatrix};\qquad \lambda=-7:\begin{bmatrix}1/\sqrt{5}\\ 2/\sqrt{5}\end{bmatrix}$$

这些特征向量自动正交（因为它们属于不同的特征值）且构成 $\mathbb{R}^2$ 的一个单位正交基. 取

$$P=\begin{bmatrix}2/\sqrt{5} & 1/\sqrt{5}\\ -1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix},\ D=\begin{bmatrix}3 & 0\\ 0 & -7\end{bmatrix}$$

那么 $A=PDP^{-1}$，且 $D=P^{-1}AP=P^{\mathrm{T}}AP$，像前面指出的那样，一个适当的变换是

$$\boldsymbol{x}=P\boldsymbol{y}\text{，此处 }\boldsymbol{x}=\begin{bmatrix}x_1\\ x_2\end{bmatrix},\quad \boldsymbol{y}=\begin{bmatrix}y_1\\ y_2\end{bmatrix}$$

那么

$$\begin{aligned}x_1^2-8x_1x_2-5x_2^2&=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}\\&=(P\boldsymbol{y})^{\mathrm{T}}A(P\boldsymbol{y})\\&=\boldsymbol{y}^{\mathrm{T}}P^{\mathrm{T}}AP\boldsymbol{y}=\boldsymbol{y}^{\mathrm{T}}D\boldsymbol{y}\\&=3y_1^2-7y_2^2\end{aligned}$$

■

为了说明例 4 中二次型相等的意义，我们可以利用新二次型计算 $Q(\boldsymbol{x})$ 在 $\boldsymbol{x}=(2,-2)$ 处的值，首先，由于 $\boldsymbol{x}=P\boldsymbol{y}$，我们得到

$$\boldsymbol{y}=P^{-1}\boldsymbol{x}=P^{\mathrm{T}}\boldsymbol{x}$$

则有

$$\boldsymbol{y}=\begin{bmatrix}2/\sqrt{5} & -1/\sqrt{5}\\ 1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}\begin{bmatrix}2\\ -2\end{bmatrix}=\begin{bmatrix}6/\sqrt{5}\\ -2/\sqrt{5}\end{bmatrix}$$

因此

$$\begin{aligned}3y_1^2-7y_2^2&=3(6/\sqrt{5})^2-7(-2/\sqrt{5})^2=3(36/5)-7(4/5)\\&=80/5=16\end{aligned}$$

这就是例 3 中 $Q(\boldsymbol{x})$ 在 $\boldsymbol{x}=(2,-2)$ 处的值，见图 7-2.

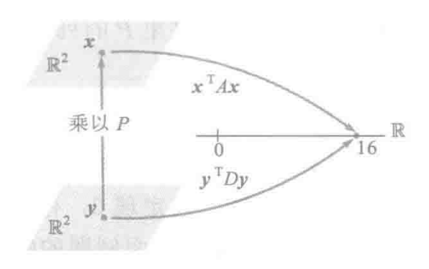


图 7-2 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 中的变量代换

例 4 说明了以下定理，定理的证明在例 4 之前已基本上给出.

定理 4 （主轴定理）

设 A 是一个 $n\times n$ 对称矩阵，那么存在一个正交变量变换 $\boldsymbol{x}=P\boldsymbol{y}$，它将二次型 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 变换为不含交叉项的二次型 $\boldsymbol{y}^{\mathrm{T}}D\boldsymbol{y}$.

定理中矩阵 P 的列称为二次型 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的**主轴**，向量 $\boldsymbol{y}$ 是向量 $\boldsymbol{x}$ 在由这些主轴构造的 $\mathbb{R}^n$ 空间的单位正交基下的坐标向量.

主轴的几何意义

若 $Q(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$，此处 A 是一个 2×2 可逆对称矩阵，c 是一个常数，可以证明 $\mathbb{R}^2$ 中所有满足

$$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=c \tag{3}$$

的 $\boldsymbol{x}$ 的集合，对应一个椭圆（或者圆）、双曲线、两条相交直线或单个点，或根本不含任意点. 如果 A 是一个对角矩阵，如图 7-3 所示，（3）的图像是标准位置. 如果 A 不是一个对角矩阵，如图 7-4 所示，（3）的图形是标准位置的旋转. 找到主轴（由 A 的特征向量确定）等同于找到一个新的坐标系统，在该坐标系统下其图形是在标准位置下的图形.

图 7-4b 中的双曲线是方程 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=16$ 的图形表示，其中 A 是例 4 中的矩阵，图 7-4b 中 y_1 轴的正向是例 4 中矩阵 P 第一列的方向，y_2 轴的正向是矩阵 P 第二列的方向.

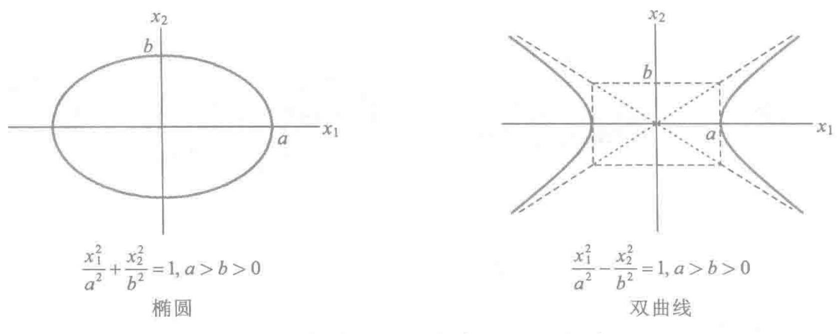


图 7-3 标准位置上的椭圆和双曲线

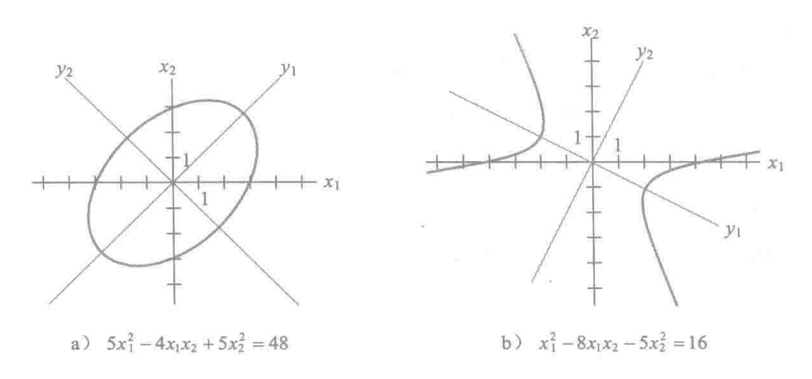


图 7-4 不在标准位置的椭圆和双曲线

例 5 图 7-4a 中，椭圆的方程为 $5x_1^2-4x_1x_2+5x_2^2=48$，求一个变量代换，将方程中的交叉项消去.

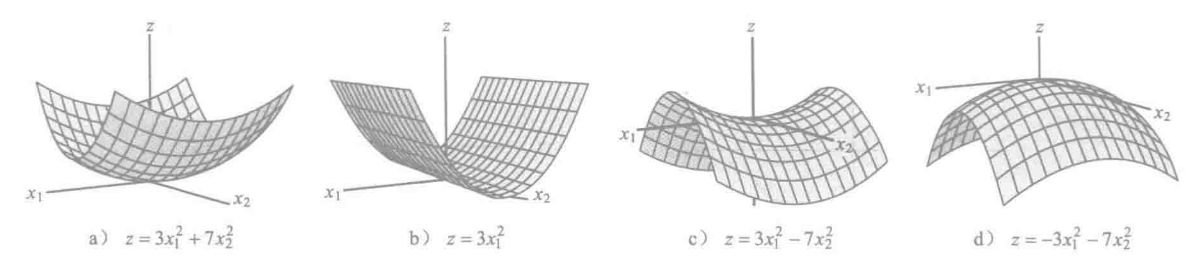
解 二次型对应的矩阵为 $A=\begin{bmatrix}5 & -2\\-2 & 5\end{bmatrix}$，$A$ 的特征值是 3 和 7，对应的单位特征向量为

$$\boldsymbol{u}_1=\begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix},\ \boldsymbol{u}_2=\begin{bmatrix}-1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}$$

令 $P=[\boldsymbol{u}_1\ \ \boldsymbol{u}_2]=\begin{bmatrix}1/\sqrt{2} & -1/\sqrt{2}\\1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}$，那么 P 可将 A 正交对角化，所以变量代换 $X=PY$ 得到的二次型为 $\boldsymbol{y}^{\mathrm{T}}D\boldsymbol{y}=3y_1^2+7y_2^2$，变量代换的新坐标轴如图 7-4a 所示. ■

二次型的分类

当 A 是一个 $n\times n$ 矩阵时，二次型 $Q(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 是一个定义域为 $\mathbb{R}^n$ 的实值函数，对二次型 $Q(\boldsymbol{x})$ 定义域中的每一个点对应 $\boldsymbol{x}=(x_1,x_2)$，可画出点 (x_1,x_2,z)，其中 $z=Q(\boldsymbol{x})$. 注意，除了 $\boldsymbol{x}=\boldsymbol{0}$ 外，图 7-5a 中所有 $Q(\boldsymbol{x})$ 的值都是正数，图 7-5d 中所有 $Q(\boldsymbol{x})$ 的值都是负值，图 7-5a 和图 7-5d 中图形的水平截面是椭圆，而图 7-5c 的水平截面对应双曲线.



a) $z=3x_1^2+7x_2^2$　b) $z=3x_1^2$　c) $z=3x_1^2-7x_2^2$　d) $z=-3x_1^2-7x_2^2$

图 7-5 二次型对应的图形

图 7-5 中简单的 2×2 例子说明下列定义.

定义 一个二次型 Q 是:

a. **正定的**，如果对所有 $\boldsymbol{x}\neq\boldsymbol{0}$，有 $Q(\boldsymbol{x})>0$.

b. **负定的**，如果对所有 $\boldsymbol{x}\neq\boldsymbol{0}$，有 $Q(\boldsymbol{x})<0$.

c. **不定的**，如果 $Q(\boldsymbol{x})$ 既有正值又有负值.

此外，Q 被称为半正定的，如果对所有 $\boldsymbol{x}$，$Q(\boldsymbol{x})\geqslant 0$；$Q$ 被称为半负定的，如果对所有 $\boldsymbol{x}$，$Q(\boldsymbol{x})\leqslant 0$. 图 7-5 中 a 和 b 的二次型都是半正定的.

定理 5 根据特征值为二次型分类. 见图 7-6.

定理 5 （二次型与特征值）

设 A 是 $n\times n$ 对称矩阵，那么一个二次型是:

a. 正定的，当且仅当 A 的所有特征值是正数.

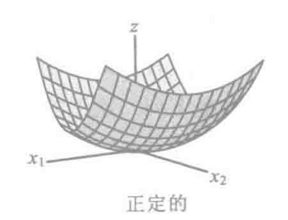
b. 负定的，当且仅当 A 的所有特征值是负数.

c. 不定的，当且仅当 A 既有正特征值，又有负特征值.

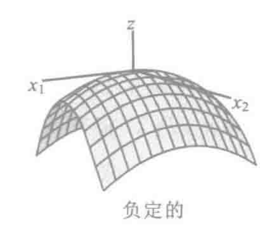
证 由主轴定理，存在一个正交变换 $\boldsymbol{x}=P\boldsymbol{y}$，使得

$$Q(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=\boldsymbol{y}^{\mathrm{T}}D\boldsymbol{y}=\lambda_1 y_1^2+\lambda_2 y_2^2+\cdots+\lambda_n y_n^2 \qquad (4)$$

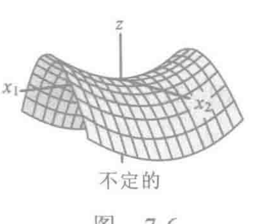
此处 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 A 的特征值，由于 P 是可逆的，非零向量 $\boldsymbol{x}$ 和非零向量 $\boldsymbol{y}$ 之间存在一个一一映射，这样，$\boldsymbol{x}\neq\boldsymbol{0}$ 时，$Q(\boldsymbol{x})$ 的值与（4）右边表达式的值完全对应. 显然，像定理所描述的三类方式一样，它由特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 的符号所确定. ■



正定的



负定的



不定的

图 7-6

例 6 $Q(\boldsymbol{x})=3x_1^2+2x_2^2+x_3^2+4x_1x_2+4x_2x_3$ 是正定的吗？

解 由于所有项的系数是正数，二次型表面上看是正定的. 但二次型对应的矩阵为

$$A=\begin{bmatrix}3&2&0\\2&2&2\\0&2&1\end{bmatrix}$$

且 A 的特征值是 5，2 和−1，所以 $Q(\boldsymbol{x})$ 是不定二次型，而不是正定二次型. ■

利用二次型的分类，相应地得到矩阵的形式分类. 一个**正定矩阵** A 是一个对称矩阵，则二次型 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 是正定的. 其他形式的矩阵（如**半正定矩阵**）的概念可类似定义.

数值计算的注解 确定对称矩阵 A 是正定的最快捷的方式，是尝试将矩阵 A 分解为 $A=R^{\mathrm{T}}R$，此处 R 是具有正对角线元素的上三角阵（可以采用一个略微修改的 LU 分解算法）. 这样的乔累斯基分解算法可行充分必要条件是 A 是正定的. 见第 7 章补充练习 7.

练习题

根据矩阵 A 的特征值，给出一个半正定矩阵.

习题 7.2

1. 计算二次型 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$，其中 $A=\begin{bmatrix}5&1/3\\1/3&1\end{bmatrix}$.

   a. $\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\end{bmatrix}$ b. $\boldsymbol{x}=\begin{bmatrix}6\\1\end{bmatrix}$ c. $\boldsymbol{x}=\begin{bmatrix}1\\3\end{bmatrix}$

2. 计算二次型 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$，其中 $A=\begin{bmatrix}4&3&0\\3&2&1\\0&1&1\end{bmatrix}$.

   a. $\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$ b. $\boldsymbol{x}=\begin{bmatrix}2\\-1\\5\end{bmatrix}$ c. $\boldsymbol{x}=\begin{bmatrix}1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{bmatrix}$

3. 求二次型的矩阵，假设 $\boldsymbol{x}$ 属于 $\mathbb{R}^2$.
   a. $10x_1^2-6x_1x_2-3x_2^2$ b. $5x_1^2+3x_1x_2$

4. 求二次型的矩阵，假设 $\boldsymbol{x}$ 属于 $\mathbb{R}^2$.
   a. $20x_1^2+15x_1x_2-10x_2^2$ b. x_1x_2

5. 求二次型的矩阵，假设 $\boldsymbol{x}$ 属于 $\mathbb{R}^3$.
   a. $8x_1^2+7x_2^2-3x_3^2-6x_1x_2+4x_1x_3-2x_2x_3$
   b. $4x_1x_2+6x_1x_3-8x_2x_3$

6. 求二次型的矩阵，假设 $\boldsymbol{x}$ 属于 $\mathbb{R}^3$.
   a. $5x_1^2-x_2^2+7x_3^2+5x_1x_2-3x_1x_3$
   b. $x_3^2-4x_1x_2+4x_2x_3$

7. 求一个变量代换，$x = Py$，将二次型 $x_1^2 + 10x_1x_2 + x_2^2$ 变换为没有交叉项的形式，给出 P 和新的二次型.

8. 设 A 是下列二次型的矩阵：

$$9x_1^2 + 7x_2^2 + 11x_3^2 - 8x_1x_2 + 8x_1x_3$$

可以证明 A 的特征是 3，9 和 15，求正交矩阵 P，使得变量代换 $x = Py$ 将 x^TAx 变换为没有交叉项的形式，给出 P 和新的二次型.

给出习题 9~18 的二次型分类，然后求一个变量代换 $x = Py$，将二次型变换成没有交叉项的形式，写出新的二次型. 利用 7.1 节的方法构造 P.

9. $3x_1^2 - 4x_1x_2 + 6x_2^2$

10. $9x_1^2 - 8x_1x_2 + 3x_2^2$

11. $2x_1^2 + 10x_1x_2 + 2x_2^2$

12. $-5x_1^2 + 4x_1x_2 - 2x_2^2$

13. $x_1^2 - 6x_1x_2 + 9x_2^2$

14. $8x_1^2 + 6x_1x_2$

15. [M] $-2x_1^2 - 6x_2^2 - 9x_3^2 - 9x_4^2 + 4x_1x_2 + 4x_1x_3 + 4x_1x_4 + 6x_3x_4$

16. [M] $4x_1^2 + 4x_2^2 + 4x_3^2 + 4x_4^2 + 3x_1x_2 + 3x_3x_4 - 4x_1x_4 + 4x_2x_3$

17. [M] $x_1^2 + x_2^2 + x_3^2 + x_4^2 + 9x_1x_2 - 12x_1x_4 + 12x_2x_3 + 9x_3x_4$

18. [M] $11x_1^2 - x_2^2 - 12x_1x_2 - 12x_1x_3 - 12x_1x_4 - 2x_3x_4$

19. 如果 $x = (x_1, x_2)$ 和 $x^Tx = 1$，即 $x_1^2 + x_2^2 = 1$，二次型 $5x_1^2 + 8x_2^2$ 的最大值是什么？（用 x 的一些例子试试.）

20. 如果 $x^Tx = 1$，二次型 $5x_1^2 - 3x_2^2$ 的最大值是什么？

习题 21~22 中，矩阵是 $n \times n$ 方阵且向量属于 $\mathbb{R}^n$，判断每个命题的真假，验证你的结论.

21. a. 二次型的矩阵是一个对称矩阵.
    b. 一个二次型没有交叉项的充分必要条件是二次型对应的矩阵是对角矩阵.
    c. 二次型 x^TAx 的主轴是 A 的特征向量.
    d. 一个正定二次型 Q 满足，对所有 $x \in \mathbb{R}^n, Q(x) > 0$.
    e. 如果一个对称矩阵 A 的所有特征值是正的，那么二次型 x^TAx 是正定的.
    f. 一个对称矩阵 A 的乔累斯基分解具有形式 $A = R^TR$，其中上三角矩阵 R 具对角线元素为正数.

22. a. 表达式 $\|x\|^2$ 是一个二次型.
    b. 如果 A 对称且 P 是正交矩阵，那么变量代换 $x = Py$ 将 x^TAx 变为没有交叉项的二次型.
    c. 如果 A 是一个 2×2 对称矩阵，那么满足 $x^TAx = c$（c 是常数）的 x 的集合对应的几何图形是圆，或椭圆或双曲线.
    d. 一个不定二次型，或者是半正定形式或者是半负定形式.
    e. 如果 A 对称且对 $x \neq 0$，二次形 x^TAx 仅仅有负值，那么 A 的所有特征值是负的.

习题 23~24 表明，当 $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}, \det A \neq 0$ 时，不用求出 A 的特征值就可将二次型 $Q(x) = x^TAx$ 分类.

23. 如果 λ_1 和 λ_2 是 A 的特征值，那么 A 的特征多项式可以用两种方式写出：$\det(A - \lambda I)$ 和 $(\lambda - \lambda_1)(\lambda - \lambda_2)$，利用这个结论说明 $\lambda_1 + \lambda_2 = a + d$（$a, d$ 是 A 的对角线上的元素)，且 $\lambda_1\lambda_2 = \det A$.

24. 验证下列命题：
    a. 如果 $\det A > 0$ 且 $a > 0$，则 Q 是正定的.
    b. 如果 $\det A > 0$ 且 $a < 0$，则 Q 是负定的.
    c. 如果 $\det A < 0$，则 Q 是不定的.

25. 证明：如果 B 是 $m \times n$ 矩阵，那么 B^TB 是半正定的；如果 B 是 $n \times n$ 可逆矩阵，那么 B^TB 是正定的.

26. 证明：如果 $n \times n$ 矩阵 A 是正定的，那么存在一个正定矩阵 B，使得 $A = B^TB$.（提示：写出 $A = PDP^T$ 且 $P^T = P^{-1}$，构造一个对角矩阵 C，使得 $D = C^TC$，且令 $B = PCP^T$，说明 B 满足条件.）

27. 令 A 和 B 是 $n \times n$ 对称矩阵且所有特征值都为正，证明 $A + B$ 的特征值也是正的.（提示：考虑二次型.）

28. 令 A 是 $n\times n$ 可逆对称矩阵，证明：如果二次型 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 是正定的，那么二次型 $\boldsymbol{x}^{\mathrm{T}}A^{-1}\boldsymbol{x}$ 也是正定的.（提示：考虑特征值.）

练习题答案

构造一个正交变量代换 $\boldsymbol{x}=P\boldsymbol{y}$，并且像（4）式一样写出：

$$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=\boldsymbol{y}^{\mathrm{T}}D\boldsymbol{y}=\lambda_1 y_1^2+\lambda_2 y_2^2+\cdots+\lambda_n y_n^2$$

如果一个特征值（例如 λ_i）是负数，那么对应于 $\boldsymbol{y}=\boldsymbol{e}_i$（$I_n$ 的第 i 列）的 $\boldsymbol{x}$ 取值使得 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 是负数，所以，一个半正定的二次型的所有特征值一定非负. 反之，如果所有特征值非负，上面的展开式说明 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 一定是半正定的. 见图 7-7.

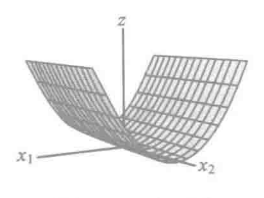


图 7-7 半正定

7.3 条件优化

工程师、经济学家、科学家和数学家常常要寻找在一些特定集合内的 $\boldsymbol{x}$ 值，使得二次型 $Q(\boldsymbol{x})$ 取最大值或最小值. 具有代表性的是，这类问题可化为 $\boldsymbol{x}$ 是在一组单位向量中的变量的优化问题. 下面我们将看到，这类条件优化问题有一个有趣且精彩的解，例 6 及后面的 7.5 节将讨论如何从实际引出这类问题.

$\mathbb{R}^n$ 中的一个单位向量 $\boldsymbol{x}$ 可用以下几种等价形式来描述：

$$\|\boldsymbol{x}\|=1\,,\ \|\boldsymbol{x}\|^2=1\,,\ \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$$

和

$$x_1^2+x_2^2+\cdots+x_n^2=1 \tag{1}$$

我们将用 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$，但在应用中经常使用展开式（1）.

当一个二次型没有交叉项时，可以很容易得到在 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 的条件下，$Q(\boldsymbol{x})$ 的最大值和最小值.

例 1 求 $Q(\boldsymbol{x})=9x_1^2+4x_2^2+3x_3^2$，在限制条件 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 下的最大值和最小值.

解 由于 x_2^2 和 x_3^2 是非负的，注意到

$$4x_2^2\leqslant 9x_2^2 \quad 和 \quad 3x_3^2\leqslant 9x_3^2$$

所以当 $x_1^2+x_2^2+x_3^2=1$ 时，

$$\begin{aligned}Q(\boldsymbol{x})&=9x_1^2+4x_2^2+3x_3^2\\&\leqslant 9x_1^2+9x_2^2+9x_3^2\\&=9(x_1^2+x_2^2+x_3^2)\\&=9\end{aligned}$$

因此当 $\boldsymbol{x}$ 为单位向量时，$Q(\boldsymbol{x})$ 的最大值不超过 9. 更进一步，当 $\boldsymbol{x}=(1,0,0)$，$Q(\boldsymbol{x})=9$. 从而 9 是 $Q(\boldsymbol{x})$ 在 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 条件下的最大值.

为求出 $Q(\boldsymbol{x})$ 的最小值，注意到

$$9x_1^2 \geqslant 3x_1^2 \text{ , } 4x_2^2 \geqslant 3x_2^2$$

因此当 $x_1^2+x_2^2+x_3^2=1$ 时，

$$Q(\boldsymbol{x}) \geqslant 3x_1^2+3x_2^2+3x_3^2=3(x_1^2+x_2^2+x_3^2)=3$$

当 $x_1=0, x_2=0, x_3=1$ 时，$Q(\boldsymbol{x})=3$，从而 3 是 $Q(\boldsymbol{x})$ 在 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 条件下的最小值. ■

从例 1 可以看到，二次型的矩阵具有特征值 9，4 和 3，且最大和最小特征值分别等于在限制条件下的 $Q(\boldsymbol{x})$ 的最大值和最小值. 我们将会看到，此结论对任何二次型也是成立的.

例 2 令 $A=\begin{bmatrix}3 & 0\\0 & 7\end{bmatrix}$，当 $\boldsymbol{x}$ 属于 $\mathbb{R}^2$ 时，$Q(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$，图 7-8 是 Q 的图形表示，图 7-9 表示圆柱体内部的一部分，圆柱与曲面的截面是点集 (x_1,x_2,z)，表示在 $x_1^2+x_2^2=1$ 情况下的 $z=Q(x_1,x_2)$. 这些点的高度值是 $Q(\boldsymbol{x})$ 的约束值，从几何意义上看，条件优化问题确定的是截面曲线上最高点和最低点的位置.

曲线上的两个最高点在 x_1x_2 平面之上 7 个单位，出现在点 $x_1=0$ 和 $x_2=\pm1$ 处，这些点对应 A 的特征值 7 和特征向量 $\boldsymbol{x}=(0,1)$ 和 $-\boldsymbol{x}=(0,-1)$. 类似地，曲线上的两个最低点在 x_1x_2 平面之上 3 个单位，它们对应特征值 3 和特征向量 (1, 0)和 (−1,0).

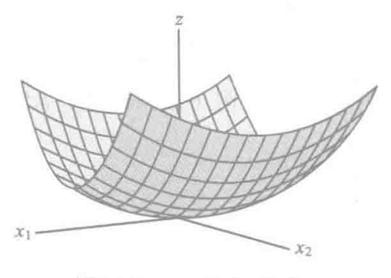


图 7-8　$z=3x_1^2+7x_2^2$

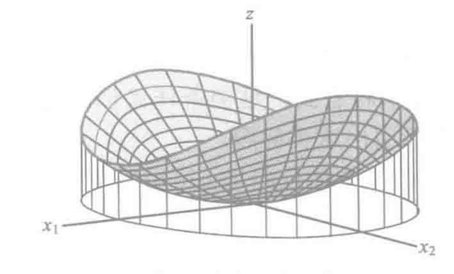


图 7-9　$z=3x_1^2+7x_2^2$ 和圆柱 $x_1^2+x_2^2=1$ 的交线 ■

图 7-9 中交线上的每一点对应的 z 坐标在 3 和 7 之间，且对任何 3 和 7 之间的数 t，存在一个单位向量 x 使得 $Q(\boldsymbol{x})=t$，换言之，所有 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的可能值在 $\|\boldsymbol{x}\|=1$ 条件下的集合是闭区间 $3 \leqslant t \leqslant 7$.

可以证明，对任何对称矩阵 A，在 $\|\boldsymbol{x}\|=1$ 条件下，$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 所有可能值的集合是实轴上的闭区间（见习题 13）. 分别用 m 和 M 表示区间的左端点和右端点，即取

$$m=\min\{\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}:\|\boldsymbol{x}\|=1\}\text{ , }M=\max\{\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}:\|\boldsymbol{x}\|=1\} \tag{2}$$

习题 12 要求你证明，如果 λ 是 A 的一个特征值，那么 $m \leqslant \lambda \leqslant M$，正像例 2 一样，下面的定理说明 m 和 M 自身也是 A 的特征值.[㊀]

㊀ (2) 式中的 min 和 max 以及定理 6 中的最小和最大，是指实数的自然顺序而不是指量级的最小和最大.

定理 6 设 A 是对称矩阵，且 m 和 M 的定义如(2)式所示，那么 M 是 A 的最大特征值 λ_1，m 是 A 的最小特征值，如果 $\boldsymbol{x}$ 是对应 M 的单位特征向量 $\boldsymbol{u}_1$，那么 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的值等于 M，如果 $\boldsymbol{x}$ 是对应 m 的单位特征向量，$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的值等于 m.

证 A 的正交对角化是 PDP^{-1}，我们知道

$$\text{当 } \boldsymbol{x} = P\boldsymbol{y} \text{ 时，} \quad \boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} = \boldsymbol{y}^{\mathrm{T}}D\boldsymbol{y} \tag{3}$$

同样

$$\text{对所有 } \boldsymbol{y}\text{，} \quad \|\boldsymbol{x}\| = \|P\boldsymbol{y}\| = \|\boldsymbol{y}\|$$

因为 $P^{\mathrm{T}}P = I$，且 $\|P\boldsymbol{y}\|^2 = (P\boldsymbol{y})^{\mathrm{T}}P\boldsymbol{y} = \boldsymbol{y}^{\mathrm{T}}P^{\mathrm{T}}P\boldsymbol{y} = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{y} = \|\boldsymbol{y}\|^2$. 特别地，$\|\boldsymbol{y}\| = 1$ 充分必要条件是 $\|\boldsymbol{x}\| = 1$，这样，如果 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 是任意单位向量，$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 和 $\boldsymbol{y}^{\mathrm{T}}D\boldsymbol{y}$ 可以认为是具有同样值的集合.

为简化记号，我们假设 A 是以 $a \geqslant b \geqslant c$ 为特征值的 3×3 矩阵，重排 P 的列（特征向量），使得 $P = [\boldsymbol{u}_1 \ \ \boldsymbol{u}_2 \ \ \boldsymbol{u}_3]$ 和

$$D = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

对给定 $\mathbb{R}^3$ 中的单位向量 $\boldsymbol{y}$，其坐标为 y_1, y_2, y_3，注意到

$$\begin{aligned} ay_1^2 &= ay_1^2 \\ by_2^2 &\leqslant ay_2^2 \\ cy_3^2 &\leqslant ay_3^2 \end{aligned}$$

将不等式相加，我们有

$$\begin{aligned} \boldsymbol{y}^{\mathrm{T}}D\boldsymbol{y} &= ay_1^2 + by_2^2 + cy_3^2 \\ &\leqslant ay_1^2 + ay_2^2 + ay_3^2 \\ &= a(y_1^2 + y_2^2 + y_3^2) \\ &= a\|\boldsymbol{y}\|^2 = a \end{aligned}$$

这样由 M 的定义可知，$M \leqslant a$. 然而，当 $\boldsymbol{y} = \boldsymbol{e}_1 = (1,0,0)$ 时，$\boldsymbol{y}^{\mathrm{T}}D\boldsymbol{y} = a$. 所以，事实上 $M = a$，且由（3）可知，对应于 $\boldsymbol{y} = \boldsymbol{e}_1$ 的向量 $\boldsymbol{x}$ 是矩阵 A 的特征向量 $\boldsymbol{u}_1$，其原因是

$$\boldsymbol{x} = P\boldsymbol{e}_1 = [\boldsymbol{u}_1 \ \boldsymbol{u}_2 \ \boldsymbol{u}_3]\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \boldsymbol{u}_1$$

这样 $M = a = \boldsymbol{e}_1^{\mathrm{T}}D\boldsymbol{e}_1 = \boldsymbol{u}_1^{\mathrm{T}}A\boldsymbol{u}_1$，从而证明了关于 M 的论断. 同理可证明 m 是最小特征值即 c 值，且 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的这个值当 $\boldsymbol{x} = P\boldsymbol{e}_3 = \boldsymbol{u}_3$ 时可以取到. ■

例 3 令 $A = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{bmatrix}$，求二次型 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 在限制条件 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} = 1$ 下的最大值，且求一个可以取到该最大值的单位向量.

解 由定理 6，我们只需求出 A 的最大特征值，其特征多项式是：

$$0=-\lambda^3+10\lambda^2-27\lambda+18=-(\lambda-6)(\lambda-3)(\lambda-1)$$

最大特征值为 6.

限制条件下 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的最大值可以在特征值 $\lambda=6$ 对应的单位特征向量 $\boldsymbol{x}$ 处获得，解 $(A-6I)\ \boldsymbol{x}=\boldsymbol{0}$，我们可得特征向量 $\begin{bmatrix}1\\1\\1\end{bmatrix}$ 和 $\boldsymbol{u}_1=\begin{bmatrix}1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{bmatrix}$. ■

在定理 7 和稍后的应用实例中，我们要考虑 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的值，这里 $\boldsymbol{x}$ 不仅是单位向量，而且是与定理 6 提到的特征向量 $\boldsymbol{u}_1$ 正交的向量，这种情形有如下的定理.

定理 7 设 A,λ_1 和 u_1 如定理 6 所示. 在如下条件限制下:

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1,\ \boldsymbol{x}^{\mathrm{T}}\boldsymbol{u}_1=0$$

$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的最大值是第二大特征值 λ_2，且这个最大值可以在 $\boldsymbol{x}$ 是对应 λ_2 的特征向量 $\boldsymbol{u}_2$ 处达到.

定理 7 的证明和上面讨论的类似，即定理化为二次型的矩阵是对角矩阵的情形，下面的例子给出对角矩阵情形下证明的思路.

例 4 求 $9x_1^2+4x_2^2+3x_3^2$ 的最大值，其限制条件是 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 和 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{u}_1=0$，这里 $\boldsymbol{u}_1=(1,0,0)$. 注意到 $\boldsymbol{u}_1$ 是二次型对应的矩阵的最大特征值 $\lambda=9$ 对应的单位特征向量.

解 如果 $\boldsymbol{x}$ 的坐标为 x_1,x_2,x_3，那么限制 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{u}_1=0$ 简单意味着 $x_1=0$，对这样的一个单位向量，$x_2^2+x_3^2=1$ 且

$$\begin{aligned}9x_1^2+4x_2^2+3x_3^2&=4x_2^2+3x_3^2\\&\leqslant 4x_2^2+4x_3^2\\&=4(x_2^2+x_3^2)\\&=4\end{aligned}$$

在这样的限制条件下，二次型的最大值不超过 4，且这个最大值可在 $\boldsymbol{x}=(0,1,0)$ 处达到，而这就是二次型对应矩阵的第二大特征值对应的特征向量. ■

例 5 令 A 表示例 3 中的矩阵，且 $\boldsymbol{u}_1$ 是对应矩阵 A 最大特征值的特征向量，求 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的最大值，其限制条件是

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1,\quad \boldsymbol{x}^{\mathrm{T}}\boldsymbol{u}_1=0 \tag{4}$$

解 由例 3 可知，A 的第二大特征值是 $\lambda=3$，解方程 $(A-3I)\boldsymbol{x}=\boldsymbol{0}$ 求出 $\lambda=3$ 对应的特征向量，且单位化得到

$$\boldsymbol{u}_2=\begin{bmatrix}1/\sqrt{6}\\1/\sqrt{6}\\-2/\sqrt{6}\end{bmatrix}$$

由于特征向量对应不同的特征值，向量 $\boldsymbol{u}_2$ 和 $\boldsymbol{u}_1$ 自然互相垂直. 这样，在条件(4)的限制下，$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的最大值是 3，且在 $\boldsymbol{x}=\boldsymbol{u}_2$ 处可以达到. ■

下面的定理是定理 7 的推广，它和定理 6 一起给出矩阵 A 所有特征值的有用特性. 具体证

明从略.

定理 8 设 A 是一个 $n\times n$ 对称矩阵，且其正交对角化为 $A=PDP^{-1}$，将对角矩阵 D 上的元素重新排列，使得 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$，且 P 的列是其对应的单位特征向量 $u_1,\cdots,u_n$. 那么对 $k=2,\cdots,n$ 时，在以下限制条件下：

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1,\ \boldsymbol{x}^{\mathrm{T}}\boldsymbol{u}_1=0,\ \cdots,\boldsymbol{x}^{\mathrm{T}}\boldsymbol{u}_{k-1}=0$$

$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的最大值是特征值 λ_k，且这个最大值在 $\boldsymbol{x}=\boldsymbol{u}_k$ 处可以达到.

定理 8 在 7.4 节和 7.5 节非常有用，下面的应用仅仅需要定理 6.

例 6 在下一年度，一个县政府计划修 x 百英里的公路和桥梁，并且修整 y 百英亩的公园和娱乐场所，政府部门必须确定在两个项目上如何分配它的资源（资金、设备和劳动等等），如果更划算的话，可以同时开始两个项目，而不是仅开始一个项目，那么 x 和 y 必须满足下面限制条件：

$$4x^2+9y^2\leqslant 36$$

见图 7-10，每个阴影可行集中的点 (x,y)，表示一个可能的该年度公共工作计划. 在限制曲线 $4x^2+9y^2=36$ 上的点，使资源利用达到最大可能.

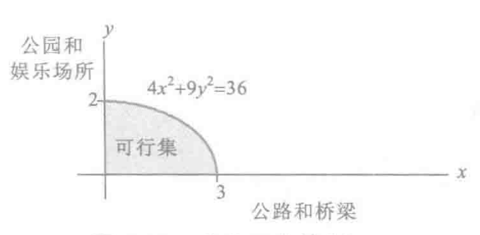


图 7-10 公共工作计划

为选择它的公共计划，县政府需要考虑居民的意见，为度量居民分配各类工作计划 (x,y) 的值或效用，经济学家有时利用下面的函数：

$$q(x,y)=xy$$

其中使 $q(x,y)$ 为常数的 (x,y) 点的集合，称为无差别曲线，从图 7-11 中可以看到三条这样的曲线，沿着无差别曲线的点对应的选择，表示居民作为一个群体有相同的价值观[㊀]. 求公共工作计划，使得效用函数 q 最大.

解 限制条件的方程 $4x^2+9y^2=36$ 并没有描述一个单位向量集，但变量代换可以修正这个问题，重写限制条件如下形式：

$$\left(\frac{x}{3}\right)^2+\left(\frac{y}{2}\right)^2=1$$

且定义

㊀ 关于无差别曲线的讨论，可参考 Michael D. Intriligator，Ronald G.Bodkin, and Cheng Hsiao, *Econometric Models, Techniques, and Application* (Upper Saddle River, NJ: Prentice-Hall, 1996).

$$x_1=\frac{x}{3},x_2=\frac{y}{2},\quad 即\ x=3x_1,y=2x_2$$

从而限制条件变成：

$$x_1^2+x_2^2=1$$

效用函数变成 $q(3x_1,2x_2)=(3x_1)(2x_2)=6x_1x_2$，取 $\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\end{bmatrix}$，那么原问题变为，在限制条件 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 下 $Q(\boldsymbol{x})=6x_1x_2$ 的最大值，注意到 $Q(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 且

$$A=\begin{bmatrix}0&3\\3&0\end{bmatrix}$$

A 的特征值是 ±3，对应 $\lambda=3$ 的特征向量是 $\begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}$，对应 $\lambda=-3$ 的特征向量是 $\begin{bmatrix}-1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}$，这样 $Q(\boldsymbol{x})=q(x_1,x_2)$ 的最大值是 3，且在 $x_1=1/\sqrt{2}$ 和 $x_2=1/\sqrt{2}$ 处可以达到.

根据原来的变量，最优的公共工作计划是修建 $x=3x_1=3/\sqrt{2}\approx2.1$ 百英里的公路和桥梁以及 $y=2x_2=\sqrt{2}\approx1.4$ 百英亩的公园和娱乐场所. 最优工作计划是限制曲线和无差别曲线 $q(x,\ y)=3$ 恰好相交的点，具有更大效用的点 (x,y) 位于和限制曲线不相交的无差别曲线上，见图 7-11.

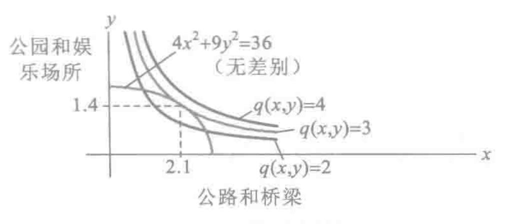


图 7-11　最优公共工作计划是（2.1,1.4）

■

练习题

1. 设 $Q(\boldsymbol{x})=3x_1^2+3x_2^2+2x_1x_2$，求一个变量代换，将 Q 变换成一个不含交叉项的二次型，且给出新二次型.
2. 对练习题 1 中的 Q，求 $Q(\boldsymbol{x})$ 在限制条件 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 下的最大值，且求出达到最大值的单位向量.

习题 7.3

在习题 1 和习题 2 中，求变量代换 $\boldsymbol{x}=P\boldsymbol{y}$，将二次型 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 变换为对应的 $\boldsymbol{y}^{\mathrm{T}}D\boldsymbol{y}$.

1. $5x_1^2+6x_2^2+7x_3^2+4x_1x_2-4x_2x_3=9y_1^2+6y_2^2+3y_3^2$
2. $3x_1^2+2x_2^2+2x_3^2+2x_1x_2+2x_1x_3+4x_2x_3=5y_1^2+2y_2^2$（提示：$\boldsymbol{x}$ 和 $\boldsymbol{y}$ 必须有同样数目的坐标，这样，这里显示的二次型关于 y_3^2 的系数一定为零.）

对习题 3~6，求（a）在条件 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 限制下，$Q(\boldsymbol{x})$ 的最大值，（b）达到最大值的一个单位向量 $\boldsymbol{u}$,（c）在条件 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 和 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{u}=0$ 限制下，$Q(\boldsymbol{x})$ 的最大值.

3. $Q(\boldsymbol{x})=5x_1^2+6x_2^2+7x_3^2+4x_1x_2-4x_2x_3$（见习题 1）
4. $Q(\boldsymbol{x})=3x_1^2+2x_2^2+2x_3^2+2x_1x_2+2x_1x_3+4x_2x_3$（见习题 2）
5. $Q(\boldsymbol{x})=5x_1^2+5x_2^2-4x_1x_2$
6. $Q(\boldsymbol{x})=7x_1^2+3x_2^2+3x_1x_2$
7. 设 $Q(\boldsymbol{x})=-2x_1^2-x_2^2+4x_1x_2+4x_2x_3$，在条件 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 的限制下，求出在 $\mathbb{R}^3$ 中的单位向量 $\boldsymbol{x}$

使 $Q(\boldsymbol{x})$ 取最大值.（提示：二次型 Q 对应矩阵的特征值是 2，-1 和-4.）

8. 设 $Q(\boldsymbol{x})=7x_1^2+x_2^2+7x_3^2-8x_1x_2-4x_1x_3-8x_2x_3$，在条件 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 的限制下，求出 $\mathbb{R}^3$ 中的单位向量 $\boldsymbol{x}$，使得 $Q(\boldsymbol{x})$ 最大.（提示：二次型 Q 对应矩阵的特征值是 9 和-3.）
9. 在条件 $x_1^2+x_2^2=1$ 的限制下，求出 $Q(\boldsymbol{x})=7x_1^2+3x_2^2-2x_1x_2$ 的最大值.（不必求出达到最大值的向量.）
10. 在条件 $x_1^2+x_2^2=1$ 的限制下，求出 $Q(\boldsymbol{x})=-3x_1^2+5x_2^2-2x_1x_2$ 的最大值.（不必求出达到最大值的向量.）
11. 若 $\boldsymbol{x}$ 是矩阵 A 对应特征值 3 的一个单位特征向量，$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的值是什么？
12. 设 λ 是对称矩阵 A 的一个特征值，验证本节的一个结论即 $m\leqslant\lambda\leqslant M$，此处 m 和 M 的定义在（2）式中给出.（提示：求出 $\boldsymbol{x}$ 使得 $\lambda=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$.）
13. 设 A 是 $n\times n$ 对称矩阵，M 和 m 表示二次型 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的最大值和最小值，且对应的单位特征向量是 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_n$，下面的计算表明，对任意给定的介于 M 和 m 之间的数 t，有一个单位向量 $\boldsymbol{x}$，使得 $t=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$，验证对 0 和 1 之间的数 α，$t=(1-\alpha)m+\alpha M$，然后令 $\boldsymbol{x}=\sqrt{1-\alpha}\boldsymbol{u}_n+\sqrt{\alpha}\boldsymbol{u}_1$，证明 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 且 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=t$.

[M]请依据习题 3~6 给出的说明做习题 14~17.

14. $x_1x_2+3x_1x_3+30x_1x_4+30x_2x_3+3x_2x_4+x_3x_4$
15. $3x_1x_2+5x_1x_3+7x_1x_4+7x_2x_3+5x_2x_4+3x_3x_4$
16. $4x_1^2-6x_1x_2-10x_1x_3-10x_1x_4-6x_2x_3-6x_2x_4-2x_3x_4$
17. $-6x_1^2-10x_2^2-13x_3^2-13x_4^2-4x_1x_2-4x_1x_3-4x_1x_4+6x_3x_4$

练习题答案

1. 二次型的矩阵 $A=\begin{bmatrix}3&1\\1&3\end{bmatrix}$，很容易求出特征值 4 和 2，对应的单位特征向量 $\begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}$ 和 $\begin{bmatrix}-1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}$，所以期望的变量代换是 $\boldsymbol{x}=P\boldsymbol{y}$，此处 $P=\begin{bmatrix}1/\sqrt{2}&-1/\sqrt{2}\\1/\sqrt{2}&1/\sqrt{2}\end{bmatrix}$.（一个常见的错误是忘记将向量单位化，新的二次型是 $\boldsymbol{y}^{\mathrm{T}}D\boldsymbol{y}=4y_1^2+2y_2^2$.）
2. 对单位向量 $\boldsymbol{x}$，$Q(\boldsymbol{x})$ 的最大值是 4，且取得最大值时的单位向量是 $\begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}$（一个常见的错误答案是 $\begin{bmatrix}1\\0\end{bmatrix}$，这个向量使二次型 $\boldsymbol{y}^{\mathrm{T}}D\boldsymbol{y}$ 而不是 $Q(\boldsymbol{x})$ 达到最大值）. 见图 7-12.

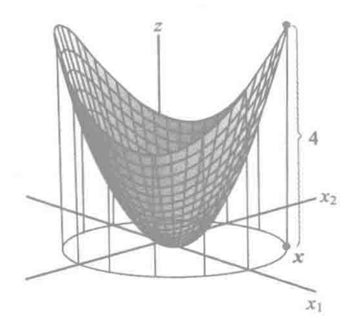


图 7-12　$Q(\boldsymbol{x})$在限制条件 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 下的最大值是 4

7.4 奇异值分解

5.3 节和 7.1 节的对角化定理在许多应用中均很重要，然而，如我们所知，不是所有矩阵都有分解式 $A=PDP^{-1}$，且 D 是对角的. 但分解 $A=QDP^{-1}$ 对任意 $m\times n$ 矩阵 A 都有可能！这类特殊分解称为奇异值分解，是线性代数应用中最有用的矩阵分解.

奇异值分解基于一般的矩阵对角化性质可以被长方形矩阵模仿：一个对称矩阵 A 的特征值的绝对值，表示度量 A 拉长或压缩一个向量（特征向量）的程度；如果 $A\boldsymbol{x}=\lambda\boldsymbol{x}$，且 $\|\boldsymbol{x}\|=1$，那么

$$\|A\boldsymbol{x}\|=\|\lambda\boldsymbol{x}\|=|\lambda|\|\boldsymbol{x}\|=|\lambda| \tag{1}$$

如果 λ_1 是具有最大数值的特征值，那么对应的单位特征向量 $\boldsymbol{v}_1$，确定一个由 A 拉长影响最大的方向，也就是说，（1）式表示当 $\boldsymbol{x}=\boldsymbol{v}_1$ 时，$A\boldsymbol{x}$ 的长度最大化，且 $\|A\boldsymbol{v}_1\|=|\lambda_1|$. 这个对 $\boldsymbol{v}_1$ 和 $|\lambda_1|$ 的描述对于长方形的矩阵来说也是类似的，这将导致奇异值分解.

例 1 如果 $A=\begin{bmatrix}4 & 11 & 14\\ 8 & 7 & -2\end{bmatrix}$，那么线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$，将 $\mathbb{R}^3$ 中的单位球 $\{\boldsymbol{x}:\|\boldsymbol{x}\|=1\}$ 映射为 $\mathbb{R}^2$ 中的椭圆，见图 7-13，找出使得长度 $\|A\boldsymbol{x}\|$ 最大的一个单位向量 $\boldsymbol{x}$，且计算这个最大长度.

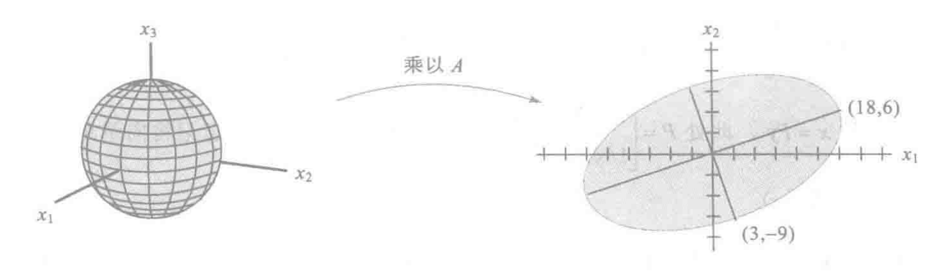


图 7-13 从 $\mathbb{R}^3$ 到 $\mathbb{R}^2$ 的一个变换

解 使 $\|A\boldsymbol{x}\|^2$ 的值最大化的 $\boldsymbol{x}$ 同样使 $\|A\boldsymbol{x}\|$ 的值最大化，但 $\|A\boldsymbol{x}\|^2$ 更容易计算，注意到

$$\|A\boldsymbol{x}\|^2=(A\boldsymbol{x})^{\mathrm{T}}(A\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}A\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}(A^{\mathrm{T}}A)\boldsymbol{x}$$

由于 $(A^{\mathrm{T}}A)^{\mathrm{T}}=A^{\mathrm{T}}A^{\mathrm{TT}}=A^{\mathrm{T}}A$，$A^{\mathrm{T}}A$ 也是一个对称矩阵. 所以，现在的问题转化为在条件 $\|\boldsymbol{x}\|=1$ 限制下，二次型 $\boldsymbol{x}^{\mathrm{T}}(A^{\mathrm{T}}A)\boldsymbol{x}$ 的最大化. 利用 7.3 节的定理 6，最大值就是矩阵 $A^{\mathrm{T}}A$ 的最大特征值 λ_1，同样，最大值可以在 $A^{\mathrm{T}}A$ 特征根 λ_1 对应的单位特征向量处获得.

对本例中的矩阵 A，

$$A^{\mathrm{T}}A=\begin{bmatrix}4 & 8\\ 11 & 7\\ 14 & -2\end{bmatrix}\begin{bmatrix}4 & 11 & 14\\ 8 & 7 & -2\end{bmatrix}=\begin{bmatrix}80 & 100 & 40\\ 100 & 170 & 140\\ 40 & 140 & 200\end{bmatrix}$$

矩阵 $A^{\mathrm{T}}A$ 的特征值是 $\lambda_1=360$，$\lambda_2=90$ 和 $\lambda_3=0$，对应的单位向量分别是

$$v_1=\begin{bmatrix}1/3\\2/3\\2/3\end{bmatrix},v_2=\begin{bmatrix}-2/3\\-1/3\\2/3\end{bmatrix},v_3=\begin{bmatrix}2/3\\-2/3\\1/3\end{bmatrix}$$

$\|Ax\|^2$ 的最大值是 360，且在 x 为单位向量 v_1 处获得，向量 Av_1 是图 7-13 中椭圆上离原点最远的点，即

$$Av_1=\begin{bmatrix}4&11&14\\8&7&-2\end{bmatrix}\begin{bmatrix}1/3\\2/3\\2/3\end{bmatrix}=\begin{bmatrix}18\\6\end{bmatrix}$$

对 $\|x\|=1$，$\|Ax\|$ 的最大值是 $\|Av_1\|=\sqrt{360}=6\sqrt{10}$. ■

例 1 说明，A 对 $\mathbb{R}^3$ 中单位球面的影响，与二次型 $x^T(A^TA)x$ 有关，事实上，变换 $x\to Ax$ 的全部几何特性都可用二次型来说明，下面我们将说明这个问题.

一个 $m\times n$ 矩阵的奇异值

令 A 是 $m\times n$ 矩阵，那么 A^TA 是对称矩阵且可以正交对角化，让 $\{v_1,\cdots,v_n\}$ 是 $\mathbb{R}^n$ 的单位正交基且构成 A^TA 的特征向量，$\lambda_1,\cdots,\lambda_n$ 是 A^TA 对应的特征值，那么对 $1\leqslant i\leqslant n$，

$$\begin{aligned}\|Av_i\|^2&=(Av_i)^TAv_i=v_i^TA^TAv_i\\&=v_i^T\ (\lambda_iv_i)\qquad\qquad \text{由于 } v_i \text{ 是 } A^TA \text{ 的特征向量}\\&=\lambda_i\qquad\qquad\qquad\quad \text{由于 } v_i \text{ 是单位向量}\end{aligned}\tag{2}$$

所以，A^TA 的所有特征值都非负，如果必要，通过重新编号我们可以假设特征值的重新排列满足

$$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n\geqslant0$$

A 的**奇异值**是 A^TA 的特征值的平方根，记为 $\sigma_1,\cdots,\sigma_n$，且它们用递减顺序排列，也就是对 $1\leqslant i\leqslant n$，$\sigma_i=\sqrt{\lambda_i}$. 由（2）可知，$A$ 的奇异值是向量 $Av_1,\cdots,Av_n$ 的长度.

例 2 令 A 是例 1 中的矩阵，由于 A^TA 的特征根是 360，90 和 0，A 的奇异值是

$$\sigma_1=\sqrt{360}=6\sqrt{10},\ \sigma_2=\sqrt{90}=3\sqrt{10},\ \sigma_3=0$$

从例 1 可知，A 的第一个奇异值是 $\|Ax\|$ 在所有单位向量处的最大值，且最大值可以在单位向量 v_1 处获得. 7.3 节的定理 7 表明，A 的第二个奇异值是，$\|Ax\|$ 在所有与 v_1 正交的单位向量上的最大值，且这个最大值可以在第二个单位特征向量 v_2 处获得（习题 22）. 对例 1 中的 v_2，

$$Av_2=\begin{bmatrix}4&11&14\\8&7&-2\end{bmatrix}\begin{bmatrix}-2/3\\-1/3\\2/3\end{bmatrix}=\begin{bmatrix}3\\-9\end{bmatrix}$$

这个点在图 7-13 中椭圆的短轴上，像 Av_1 在长轴上一样（见图 7-14），矩阵 A 的前两个奇异值是椭圆长半轴和短半轴的长度.

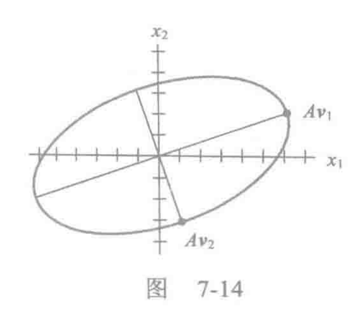


图 7-14 ■

事实上，像下面定理所说，图 7-14 中的 $A\boldsymbol{v}_1$ 和 $A\boldsymbol{v}_2$ 正交不是偶然.

定理 9 假若 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$ 是包含 $A^{\mathrm{T}}A$ 特征向量的 $\mathbb{R}^n$ 上的单位正交基，重新整理使得对应 $A^{\mathrm{T}}A$ 的特征值满足 $\lambda_1 \geqslant \cdots \geqslant \lambda_n$. 假若 A 有 r 个非零奇异值，那么 $\{A\boldsymbol{v}_1,\cdots,A\boldsymbol{v}_r\}$ 是 Col$A$ 的一个正交基，且 rank $A = r$.

证 由于当 $i \neq j$ 时，$\boldsymbol{v}_i$ 和 $\lambda_j \boldsymbol{v}_j$ 正交，所以

$$(A\boldsymbol{v}_i)^{\mathrm{T}}(A\boldsymbol{v}_j) = \boldsymbol{v}_i^{\mathrm{T}} A^{\mathrm{T}} A \boldsymbol{v}_j = \boldsymbol{v}_i^{\mathrm{T}}(\lambda_j \boldsymbol{v}_j) = 0$$

那么 $\{A\boldsymbol{v}_1,\cdots,A\boldsymbol{v}_n\}$ 是一个正交基，更进一步，由于向量 $A\boldsymbol{v}_1,\cdots,A\boldsymbol{v}_n$ 的长度是 A 的奇异值，且因为有 r 个非零奇异值，$A\boldsymbol{v}_i \neq \boldsymbol{0}$ 的充分必要条件是 $1 \leqslant i \leqslant r$. 所以 $A\boldsymbol{v}_1,\cdots,A\boldsymbol{v}_r$ 是线性无关向量，且属于 Col$A$. 最后，对任意属于 Col A 的 $\boldsymbol{y}$，如 $\boldsymbol{y} = A\boldsymbol{x}$，我们可以写出 $\boldsymbol{x} = c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n$，且

$$\begin{aligned}\boldsymbol{y} = A\boldsymbol{x} &= c_1 A\boldsymbol{v}_1 + \cdots + c_r A\boldsymbol{v}_r + c_{r+1} A\boldsymbol{v}_{r+1} + \cdots + c_n A\boldsymbol{v}_n \\ &= c_1 A\boldsymbol{v}_1 + \cdots + c_r A\boldsymbol{v}_r + 0 + \cdots + 0\end{aligned}$$

这样，$\boldsymbol{y}$ 在 Span$\{A\boldsymbol{v}_1,\cdots,A\boldsymbol{v}_r\}$ 中，这说明 $\{A\boldsymbol{v}_1,\cdots,A\boldsymbol{v}_r\}$ 是 Col A 的一个正交基. 因此 rank $A = \dim(\mathrm{Col}\ A) = r$. ■

数值计算的注解 在一些情形下，A 的秩对 A 中元素的微小变化很敏感，如果用计算机对矩阵 A 作行简化，明显矩阵 A 的主元列数目的计算方法效果不好，舍入误差常导致满秩的阶梯矩阵.

实际上，估计大矩阵的秩时，最可靠的方法是计算非零奇异值的个数，在这种情形下，特别小的非零奇异值在实际计算中常假定为零，矩阵 A 的有效秩是剩余非零奇异值的数目.㊀

奇异值分解

矩阵 A 的分解涉及一个 $m \times n$ "对角"矩阵 Σ，其形式是

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \begin{matrix} \\ \longleftarrow m-r \text{ 行} \end{matrix} \qquad (3)$$

$$\uparrow\!___ \ n-r \text{ 列}$$

此处，D 是一个 $r \times r$ 对角矩阵，且 r 不超过 m 和 n 的最小值（如果 r 等于 m 或 n，或都相等，则

㊀ 一般地，秩的估计问题并不简单，对这个问题的详细讨论见 Philip E. Gill, Walter Murray, and Margaret H. Wright, *Numerical Linear Algebra and Optimization*, vol.1 (Redwood City, CA: Addison-Wesley, 1991), 5.8 节.

M 中不会出现零矩阵）.

定理 10 （奇异值分解）

设 A 是秩为 r 的 $m\times n$ 矩阵，那么存在一个类似（3）的 $m\times n$ 矩阵 Σ，此处 D 的对角线元素是 A 的前 r 个奇异值，$\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_r>0$，并且存在一个 $m\times m$ 正交矩阵 U 和一个 $n\times n$ 正交矩阵 V 使得 $A=U\Sigma V^{\mathrm{T}}$.

任何分解 $A=U\Sigma V^{\mathrm{T}}$ 称为 A 的一个**奇异值分解**（或 SVD），其中 U 和 V 是正交矩阵，Σ形如（3）式，且 D 具有正的对角线元素，矩阵 U 和 V 不是由 A 惟一确定的，但Σ的对角线元素必须是 A 的奇异值，见习题 19，分解中 U 的列称为 A 的**左奇异向量**，而 V 的列称为 A 的**右奇异向量**.

证 假设 λ_i和$\boldsymbol{v}_j$ 如定理 9，使得 $\{A\boldsymbol{v}_1,\cdots,A\boldsymbol{v}_r\}$ 是 Col A 的正交基，将每一个 $A\boldsymbol{v}_i$ 单位化得到一个标准正交基 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_r\}$，此处

$$\boldsymbol{u}_i=\frac{1}{\|A\boldsymbol{v}_i\|}A\boldsymbol{v}_i=\frac{1}{\sigma_i}A\boldsymbol{v}_i$$

而且

$$A\boldsymbol{v}_i=\sigma_i\boldsymbol{u}_i\quad(1\leqslant i\leqslant r)\tag{4}$$

现在将 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_r\}$ 扩充为 $\mathbb{R}^m$ 的单位正交基 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_m\}$. 而且取

$$U=[\boldsymbol{u}_1\quad\boldsymbol{u}_2\quad\cdots\quad\boldsymbol{u}_m]\quad\text{和}\quad V=[\boldsymbol{v}_1\quad\boldsymbol{v}_2\quad\cdots\quad\boldsymbol{v}_n]$$

由构造可知，U和V 是正交矩阵，同样由（4）

$$AV=[A\boldsymbol{v}_1\ \cdots\ A\boldsymbol{v}_r\ \ \boldsymbol{0}\ \cdots\ \boldsymbol{0}]=[\sigma_1\boldsymbol{u}_1\ \cdots\ \sigma_r\boldsymbol{u}_r\ \ \boldsymbol{0}\ \cdots\ \boldsymbol{0}]$$

设 D 是对角元素为 $\sigma_1,\cdots,\sigma_r$ 的对角矩阵，且Σ是上面（3）中的形式，那么

$$\begin{aligned}U\Sigma&=[\boldsymbol{u}_1\ \boldsymbol{u}_2\cdots\boldsymbol{u}_m]\left[\begin{array}{cccc|c}\sigma_1&&&0&\\&\sigma_2&&&0\\&&\ddots&&\\0&&&\sigma_r&\\\hline&&0&&0\end{array}\right]\\&=[\sigma_1\boldsymbol{u}_1\cdots\sigma_r\boldsymbol{u}_r\ \boldsymbol{0}\cdots\boldsymbol{0}]\\&=AV\end{aligned}$$

由于V 是一个正交矩阵，$U\Sigma V^{\mathrm{T}}=AVV^{\mathrm{T}}=A$. ■

下面的两个例子注重于讨论奇异值分解的内部结构. 一个快速而数值稳定的分解算法将会使用另外一种方法. 参见本节末尾处的“数值计算的注解”.

例 3 使用例 1 和例 2 中的结果求 $A=\begin{bmatrix}4&11&14\\8&7&-2\end{bmatrix}$的一个奇异值分解.

解 一个奇异值分解可分三步进行.

第一步. 将矩阵 $A^{\mathrm{T}}A$ 正交对角化. 即求矩阵 $A^{\mathrm{T}}A$ 的特征值及其对应的特征向量的标准正交集. 如果 A 只有两列，其计算可以手算完成. 对规模较大的矩阵通常需要使用矩阵程序. [1] 而对

---

[1] 参阅软件和几何计算器命令的学习指南. 例如，MATLAB 用一个命令 eig 就可以求出特征值和特征向量.

此处的矩阵 A，在例 1 中已经给出了 $A^{\mathrm{T}}A$ 的特征值及其相应的特征向量.

第二步. 算出 V 和Σ. 将 $A^{\mathrm{T}}A$ 的特征值按降序排列. 在例 1 中,特征值已经按降序排列:360，90，0. 它们对应的单位特征向量分别为 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$，是 A 的右奇异向量. 使用例 1 的结果，构造

$$V=[\boldsymbol{v}_1\quad \boldsymbol{v}_2\quad \boldsymbol{v}_3]=\begin{bmatrix}1/3 & -2/3 & 2/3\\ 2/3 & -1/3 & -2/3\\ 2/3 & 2/3 & 1/3\end{bmatrix}$$

特征值的算术平方根就是奇异值：$\sigma_1=\sqrt{360}=6\sqrt{10},\ \sigma_2=\sqrt{90}=3\sqrt{10},\ \sigma_3=0$. 其中的非零奇异值就是矩阵 D 的对角元素. 矩阵Σ与矩阵 A 的行列数相同，以矩阵 D 为其左上角，其他元素为 0.

$$D=\begin{bmatrix}6\sqrt{10} & 0\\ 0 & 3\sqrt{10}\end{bmatrix},\ \Sigma=[D\quad 0]=\begin{bmatrix}6\sqrt{10} & 0 & 0\\ 0 & 3\sqrt{10} & 0\end{bmatrix}$$

第三步. 构造 U. 当矩阵 A 的秩为 r 时，矩阵 U 的前 r 列是从 $A\boldsymbol{v}_1\cdots A\boldsymbol{v}_r$ 计算得到的单位向量. 此例中，A 有两个非零奇异值，因此 $\operatorname{rank} A=2$. 根据例 2 之前的内容和方程（2），有 $\|A\boldsymbol{v}_1\|=\sigma_1$，$\|A\boldsymbol{v}_2\|=\sigma_2$，于是

$$\boldsymbol{u}_1=\frac{1}{\sigma_1}A\boldsymbol{v}_1=\frac{1}{6\sqrt{10}}\begin{bmatrix}18\\ 6\end{bmatrix}=\begin{bmatrix}3/\sqrt{10}\\ 1/\sqrt{10}\end{bmatrix}$$

$$\boldsymbol{u}_2=\frac{1}{\sigma_2}A\boldsymbol{v}_2=\frac{1}{3\sqrt{10}}\begin{bmatrix}3\\ -9\end{bmatrix}=\begin{bmatrix}1/\sqrt{10}\\ -3/\sqrt{10}\end{bmatrix}$$

留意到 $\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$ 已是 $\mathbb{R}^2$ 的一个基，因此构造 U 不需另外的向量，$U=[\boldsymbol{u}_1\quad \boldsymbol{u}_2]$. A 的奇异值分解是

$$A=\underset{\displaystyle\uparrow \atop \displaystyle U}{\begin{bmatrix}3/\sqrt{10} & 1/\sqrt{10}\\ 1/\sqrt{10} & -3/\sqrt{10}\end{bmatrix}}\ \underset{\displaystyle\uparrow \atop \displaystyle \Sigma}{\begin{bmatrix}6\sqrt{10} & 0 & 0\\ 0 & 3\sqrt{10} & 0\end{bmatrix}}\ \underset{\displaystyle\uparrow \atop \displaystyle V^{\mathrm{T}}}{\begin{bmatrix}1/3 & 2/3 & 2/3\\ -2/3 & -1/3 & 2/3\\ 2/3 & -2/3 & 1/3\end{bmatrix}}$$

■

例 4 求 $A=\begin{bmatrix}1 & -1\\ -2 & 2\\ 2 & -2\end{bmatrix}$的一个奇异值分解.

解 首先，计算 $A^{\mathrm{T}}A=\begin{bmatrix}9 & -9\\ -9 & 9\end{bmatrix}$，$A^{\mathrm{T}}A$ 的特征值是 18 和 0，相应的单位特征向量是

$$\boldsymbol{v}_1=\begin{bmatrix}1/\sqrt{2}\\ -1/\sqrt{2}\end{bmatrix},\ \boldsymbol{v}_2=\begin{bmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix}$$

这两个向量构成 V 的列向量：

$$V=[\boldsymbol{v}_1\quad \boldsymbol{v}_2]=\begin{bmatrix}-1/\sqrt{2} & 1/\sqrt{2}\\ -1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}$$

矩阵的奇异值为 $\sigma_1=\sqrt{18}=3\sqrt{2},\ \sigma_2=0$. 因为只有一个非零的奇异值，“矩阵” D 可写成单个数值，即 $D=3\sqrt{2}$. 矩阵Σ与矩阵 A 的行列数相同，以矩阵 D 为其左上角：

$$\Sigma=\begin{bmatrix}D&0\\0&0\\0&0\end{bmatrix}=\begin{bmatrix}3\sqrt{2}&0\\0&0\\0&0\end{bmatrix}$$

为构造U，首先计算$A\boldsymbol{v}_1$和$A\boldsymbol{v}_2$：

$$A\boldsymbol{v}_1=\begin{bmatrix}2/\sqrt{2}\\-4/\sqrt{2}\\4/\sqrt{2}\end{bmatrix},\quad A\boldsymbol{v}_2=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

为检查计算，验证$\|A\boldsymbol{v}_1\|=\sigma_1=\sqrt{18}=3\sqrt{2}$. 当然，$A\boldsymbol{v}_2=0$，因为$\|A\boldsymbol{v}_2\|=\sigma_2=0$. 目前只找到$U$的一列是

$$\boldsymbol{u}_1=\frac{1}{3\sqrt{2}}A\boldsymbol{v}_1=\begin{bmatrix}1/3\\-2/3\\2/3\end{bmatrix}$$

U的其他列是将集合$\{\boldsymbol{u}_1\}$扩充为$\mathbb{R}^3$的标准正交基而得到的. 此时，我们需要两个单位正交向量且都与$\boldsymbol{u}_1$正交.（见图 7-15.）每个向量必须满足$\boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{x}=0$，这等价于方程$x_1-2x_2+2x_3=0$，该方程的解构成的基是

$$\boldsymbol{w}_1=\begin{bmatrix}2\\1\\0\end{bmatrix},\ \boldsymbol{w}_2=\begin{bmatrix}-2\\0\\1\end{bmatrix}$$

（经检验$\boldsymbol{w}_1$和$\boldsymbol{w}_2$都与$\boldsymbol{u}_1$正交.）应用格拉姆-施密特正交法于$\{\boldsymbol{w}_1,\boldsymbol{w}_2\}$，可以得到

$$\boldsymbol{u}_2=\begin{bmatrix}2/\sqrt{5}\\1/\sqrt{5}\\0\end{bmatrix},\ \boldsymbol{u}_3=\begin{bmatrix}-2/\sqrt{45}\\4/\sqrt{45}\\5/\sqrt{45}\end{bmatrix}$$

最后，令$U=[\boldsymbol{u}_1\quad\boldsymbol{u}_2\quad\boldsymbol{u}_3]$,以及由上所得的$V^{\mathrm{T}}$和$\Sigma$，那么

$$A=\begin{bmatrix}1&-1\\-2&2\\2&-2\end{bmatrix}=\begin{bmatrix}1/3&2/\sqrt{5}&-2/\sqrt{45}\\-2/3&1/\sqrt{5}&4/\sqrt{45}\\2/3&0&5/\sqrt{45}\end{bmatrix}\begin{bmatrix}3\sqrt{2}&0\\0&0\\0&0\end{bmatrix}\begin{bmatrix}1/\sqrt{2}&-1/\sqrt{2}\\1/\sqrt{2}&1/\sqrt{2}\end{bmatrix}$$

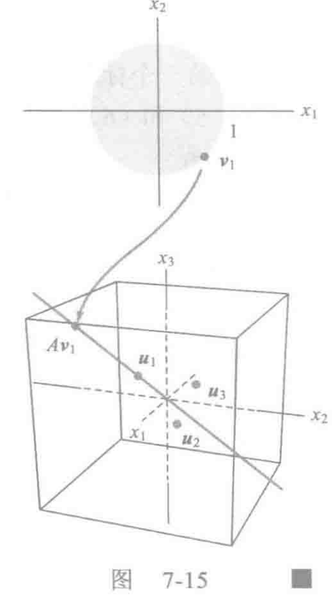


图 7-15 ■

奇异值分解的应用

如上所述，奇异值分解常用于估计矩阵的秩，某些数值计算的其他应用，下面简单叙述它在数值计算中的其他应用，7.5 节给出一个奇异值分解在图像处理的应用.

例 5（条件数） 当应用矩阵A的奇异值分解时，多数涉及方程$A\boldsymbol{x}=\boldsymbol{b}$的数值计算要尽可能的可靠. 两个正交矩阵$U$和$V$不影响向量的长度和两个向量的夹角（6.2 节定理 7）. 数值计算中的任何不稳定因素都与Σ有关. 如果A的奇异值非常大或非常小，舍入误差几乎不可避免，

此时，知道Σ和V中的元素对分析误差特别有用.

如果A是一个$n\times n$可逆矩阵，那么最大奇异值和最小奇异值的比率σ_1/σ_n，给出矩阵A的**条件数**，2.3节的习题41~43表明，条件数如何影响$A\boldsymbol{x}=\boldsymbol{b}$的解对$A$的元素变化（或误差）的敏感程度（事实上，$A$的“条件数”可用几种方式计算，但这里给出的定义可广泛用于$A\boldsymbol{x}=\boldsymbol{b}$的研究）. ■

例6（基本子空间的基） 给定$m\times n$矩阵A的一个奇异值分解，取$\boldsymbol{u}_1,\cdots,\boldsymbol{u}_m$是左奇异向量，$\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$是右奇异向量，且$\sigma_1,\cdots,\sigma_n$是奇异值，$r$为$A$的秩，由定理9，

$$\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_r\} \tag{5}$$

是Col A的一个单位正交基. 见图7-16.

回忆6.1节的定理3，$(\text{Col }A)^{\perp}=\text{Nul }A^{\mathrm{T}}$，因此

$$\{\boldsymbol{u}_{r+1},\cdots,\boldsymbol{u}_m\} \tag{6}$$

是Nul A^{T}的一个标准正交基.

由于当$1\leqslant i\leqslant n$时$\|A\boldsymbol{v}_i\|=\sigma_i$，且$\sigma_i$是零的充分必要条件是$i>r$，向量$\boldsymbol{v}_{r+1},\cdots,\boldsymbol{v}_n$生成一个维数为$(n-r)$的子空间Nul A. 由秩定理可知，$\dim\text{Nul }A=n-\text{rank }A$，从而说明

$$\{\boldsymbol{v}_{r+1},\cdots,\boldsymbol{v}_n\} \tag{7}$$

是Nul A的一个标准正交基，见4.5节的基定理.

从（5）和（6），空间Nul A^{T}的正交补是Col A，将A和A^{T}交换，我们有

$$(\text{Nul }A)^{\perp}=\text{Col }A^{\mathrm{T}}=\text{Row }A$$

因此，从（7）得

$$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r\} \tag{8}$$

是Row A的一个标准正交基.

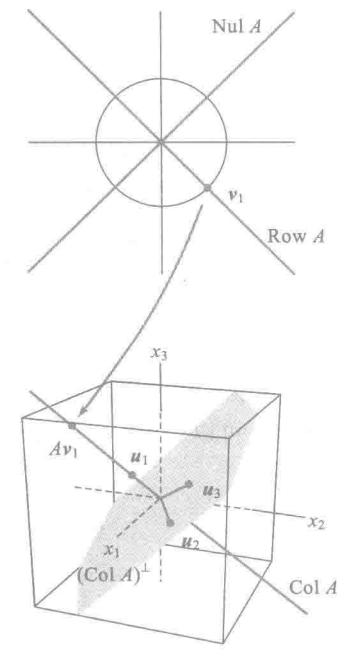


图7-16 例6中的基本子空间

图7-17给出（5）~（8）的总结，但说明的是$\{\sigma_1\boldsymbol{u}_1,\cdots,\sigma_r\boldsymbol{u}_r\}$为Col A的正交基，而不是标准基. 注意到，对$1\leqslant i\leqslant r$，$A\boldsymbol{v}_i=\sigma_i\boldsymbol{u}_i$. 由$A$确定的四个基本子空间的标准正交基在某些计算中非常有用，特别在条件优化问题中更是如此.

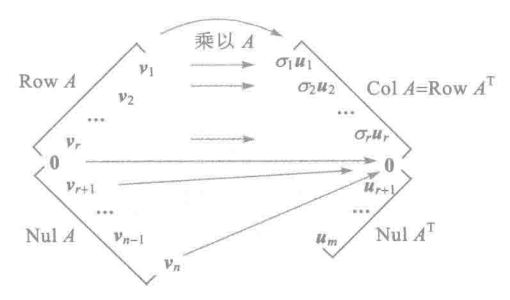


图7-17 四个基本子空间和A的作用 ■

四个基本子空间和奇异值的概念给可逆矩阵定理提供了最后的命题.（回忆为了避免几乎将可逆矩阵定理的命题数目翻倍，定理中已将与 A^{T} 有关的命题删除了.）可逆矩阵定理的其他命题在 2.3 节、2.9 节、3.2 节、4.6 节和 5.2 节中已经给出.

定理　（可逆矩阵定理（最后补充））

设 A 为 $n\times n$ 矩阵，那么下述命题中的每一个都与 A 是可逆矩阵的命题等价.

u. $(\mathrm{Col}\,A)^{\perp}=\{\mathbf{0}\}$.

v. $(\mathrm{Nul}\,A)^{\perp}=\mathbb{R}^n$.

w. $\mathrm{Row}\,A=\mathbb{R}^n$.

x. *A 有 n 个非零的奇异值.*

例 7（奇异值分解的简化和 A 的伪逆）　当Σ包含零元素的行或列时，矩阵 A 具有更简洁的分解，利用上面建立的符号，取 $r=\mathrm{rank}\,A$，将U和V 矩阵分块为第一块包含 r 列的子矩阵：

$$U=[U_r\quad U_{m-r}],\quad \text{其中}\,U_r=[\boldsymbol{u}_1\quad\cdots\quad\boldsymbol{u}_r]$$
$$V=[V_r\quad V_{n-r}],\quad \text{其中}\,V_r=[\boldsymbol{v}_1\quad\cdots\quad\boldsymbol{v}_r]$$

那么U_r是 $m\times r$ 和V_r是 $n\times r$.（为简化符号，我们考虑U_{m-r}和V_{n-r}，即使其中一个没有任何列.）分块矩阵的乘法表明

$$A=[U_r\ \ U_{m-r}]\begin{bmatrix}D & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}V_r^{\mathrm{T}}\\ V_{n-r}^{\mathrm{T}}\end{bmatrix}=U_rDV_r^{\mathrm{T}}\tag{9}$$

矩阵 A 的这个分解称为 A 的**简化的奇异值分解**. 由于 D 的对角元素非零，我们得到下列形式的矩阵，称为 A 的**伪逆**（也称**穆尔-彭罗斯逆**）：

$$A^{+}=V_rD^{-1}U_r^{\mathrm{T}}\tag{10}$$

本章最后的补充习题 12~14 研究了简化的奇异值分解和伪逆的一些性质.　■

例 8（最小二乘解）　给定方程 $A\boldsymbol{x}=\boldsymbol{b}$，利用式（10）中给出 A 的伪逆，定义

$$\hat{\boldsymbol{x}}=A^{+}\boldsymbol{b}=V_rD^{-1}U_r^{\mathrm{T}}\boldsymbol{b}$$

那么由（9）式中的奇异值分解得到

$$\begin{aligned}A\hat{\boldsymbol{x}}&=(U_rDV_r^{\mathrm{T}})(V_rD^{-1}U_r^{\mathrm{T}}\boldsymbol{b})\\&=U_rDD^{-1}U_r^{\mathrm{T}}\boldsymbol{b}&&\text{由于 }V_r^{\mathrm{T}}V_r=I_r\\&=U_rU_r^{\mathrm{T}}\boldsymbol{b}\end{aligned}$$

从（5）式可知，$U_rU_r^{\mathrm{T}}\boldsymbol{b}$ 是 $\boldsymbol{b}$ 在 $\mathrm{Col}\,A$ 上的正交投影 $\hat{\boldsymbol{b}}$（见 6.3 节定理 10），因此 $\hat{\boldsymbol{x}}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解，实际上，这个 $\hat{\boldsymbol{x}}$ 在 $A\boldsymbol{x}=\boldsymbol{b}$ 的所有最小二乘解中具有最小长度，见补充习题 14.　■

数值计算的注解　例 1~4 和习题说明了奇异值概念并给出如何手工计算，实际上，应该避免 $A^{\mathrm{T}}A$ 的计算，原因是任何 A 中元素的误差在 $A^{\mathrm{T}}A$ 中被平方. 存在快速的迭代方法，可计算精确到很多位数的矩阵 A 的奇异值和奇异向量.

进一步阅读

Horn, Roger A., and Charles R. Johnson, *Matrix Analysis*[⊖], vol. 1(Cambridge: Cambridge University Press, 1985), pp. 414–445.

Long, Cliff, "Visualization of Matrix Singular Value Decomposition." *Mathematics Magazine* **56** (1983), pp. 161–167.

Moler, C. B., and D. Morrison, "Singular Value Analysis of Cryptograms." *Amer. Math. Monthly* **90** (1983), pp. 78–87.

Strang, Gilbert, *Linear Algebra and Its Applications,* 3rd ed. (San Diego: Harcourt Brace Jovanovich, 1988), pp. 442–452.

Watkins, David S., *Fundamentals of Matrix Computations* (New York: Wiley, 1991), pp. 390–398, 409–421.

练习题

给定一个奇异值分解，$A=U\Sigma V^{\mathrm{T}}$，求 A^{T} 的一个奇异值分解，A 和 A^{T} 的奇异值有什么联系？

习题 7.4

求习题 1~4 中矩阵的奇异值分解.

1. $\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$

2. $\begin{bmatrix} -5 & 0 \\ 0 & 0 \end{bmatrix}$

3. $\begin{bmatrix} \sqrt{6} & 1 \\ 0 & \sqrt{6} \end{bmatrix}$

4. $\begin{bmatrix} \sqrt{3} & 2 \\ 0 & \sqrt{3} \end{bmatrix}$

求习题 5~12 中每个矩阵的奇异值分解.（提示：在习题 11 中，一个 U 的选择可以是 $\begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix}$，习题 12 中 U 的一个列可以是 $\begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$.）

5. $\begin{bmatrix} -3 & 0 \\ 0 & 0 \end{bmatrix}$

6. $\begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}$

7. $\begin{bmatrix} 2 & -1 \\ 2 & 2 \end{bmatrix}$

8. $\begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$

9. $\begin{bmatrix} 7 & 1 \\ 0 & 0 \\ 5 & 5 \end{bmatrix}$

10. $\begin{bmatrix} 4 & -2 \\ 2 & -1 \\ 0 & 0 \end{bmatrix}$

11. $\begin{bmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{bmatrix}$

12. $\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}$

13. 求 $A=\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$ 的奇异值分解（提示：对 A^{T} 进行分解）.

14. 在习题 7 中，求一个单位向量 $\boldsymbol{x}$，使得 $A\boldsymbol{x}$ 具有最大长度.

15. 假设下面是矩阵 A 的奇异值分解，其中 U 和 V 的元素是四舍五入精确到小数点后两位数字.

$$A=\begin{bmatrix} 0.40 & -0.78 & 0.47 \\ 0.37 & -0.33 & -0.87 \\ -0.84 & -0.52 & -0.16 \end{bmatrix}\begin{bmatrix} 7.10 & 0 & 0 \\ 0 & 3.10 & 0 \\ 0 & 0 & 0 \end{bmatrix}\times\begin{bmatrix} 0.30 & -0.51 & -0.81 \\ 0.76 & 0.64 & -0.12 \\ 0.58 & -0.58 & 0.58 \end{bmatrix}$$

a. A 的秩是多少？

b. 利用 A 的这个分解，不用计算，写出 $\operatorname{Col} A$ 的一个基和 $\operatorname{Nul} A$ 的一个基.（提示：首先写出 V 的列.）

[⊖] 该书中文版已由机械工业出版社引进出版. ——编辑注

16. 对下列 3×4 矩阵 A 的奇异值分解，重做习题 15.

$$A=\begin{bmatrix}-0.86 & -0.11 & -0.50\\ 0.31 & 0.68 & -0.67\\ 0.41 & -0.73 & -0.55\end{bmatrix}\begin{bmatrix}12.48 & 0 & 0 & 0\\ 0 & 6.34 & 0 & 0\\ 02.48 & 0 & 0 & 0\end{bmatrix}\times$$
$$\begin{bmatrix}0.66 & -0.03 & -0.35 & 0.66\\ -0.13 & -0.90 & -0.39 & -0.13\\ 0.65 & 0.08 & -0.16 & -0.73\\ -0.34 & 0.42 & -0.84 & -0.08\end{bmatrix}$$

在习题 17~24 中，A 是一个 $m\times n$ 矩阵，具有奇异值分解 $A=U\Sigma V^{\mathrm{T}}$，此处 U 是一个 $m\times m$ 正交矩阵，Σ 是一个“对角”矩阵，有 r 个正元素，没有负元素，V 是一个 $n\times n$ 正交矩阵，验证每一个答案.

17. 若 A 是可逆方阵，求 A^{-1} 的奇异值分解.
18. 证明：如果 A 是方阵，那么 $|\det A|$ 是 A 的奇异值的乘积.
19. 证明：V 的列是 $A^{\mathrm{T}}A$ 的特征向量，U 的列是 AA^{T} 的特征向量，而 Σ 对角线上的元素是 A 的奇异值.（提示：利用奇异值分解计算 $A^{\mathrm{T}}A$ 和 AA^{T}.）
20. 证明：如果 A 是 $n\times n$ 且正定的矩阵，那么 A 的正交对角化 $A=PDP^{\mathrm{T}}$ 就是 A 的奇异值分解.
21. 证明：如果 P 是一个 $m\times m$ 正交矩阵，那么 PA与A 有同样的奇异值.
22. 验证例 2 中的论断，当 $\boldsymbol{x}$ 变化范围属于所有垂直于 $\boldsymbol{v}_1$ 的单位向量时，矩阵 A 的第二个奇异值是，$\|A\boldsymbol{x}\|$ 的最大值，这里 $\boldsymbol{v}_1$ 是对应 A 的第一个奇异值的右奇异向量.（提示：利用 7.3 节的定理 7.）
23. 如果 $U=[\boldsymbol{u}_1 \ \cdots \ \boldsymbol{u}_m]$和$V=[\boldsymbol{v}_1 \ \cdots \ \boldsymbol{v}_n]$，证明：
$$A=\sigma_1\boldsymbol{u}_1\boldsymbol{v}_1^{\mathrm{T}}+\sigma_2\boldsymbol{u}_2\boldsymbol{v}_2^{\mathrm{T}}+\cdots+\sigma_r\boldsymbol{u}_r\boldsymbol{v}_r^{\mathrm{T}}$$
24. 利用习题 23 的记号，说明对 $1\leqslant j\leqslant r=\operatorname{rank} A$，$A^{\mathrm{T}}\boldsymbol{u}_j=\sigma_j\boldsymbol{v}_j$ 成立.
25. 令 $T:\mathbb{R}^n\to\mathbb{R}^m$ 是线性变换，描述如何求 $\mathbb{R}^n$ 的基 B 和 $\mathbb{R}^m$ 中的基 C，使得矩阵 T 相对于 B 和 C 是一个 $m\times n$ “对角”矩阵.

[M]计算习题 26 和习题 27 中每一个矩阵的奇异值分解，最后的矩阵元素精确到两位小数，利用例 3 和例 4 的方法.

26. $A=\begin{bmatrix}-18 & 13 & -4 & 4\\ 2 & 19 & -4 & 12\\ -14 & 11 & -12 & 8\\ -2 & 21 & 4 & 8\end{bmatrix}$

27. $A=\begin{bmatrix}6 & -8 & -4 & 5 & -4\\ 2 & 7 & -5 & -6 & 4\\ 0 & -1 & -8 & 2 & 2\\ -1 & -2 & 4 & 4 & -8\end{bmatrix}$

28. [M]计算 2.3 节习题 9 中 4×4 矩阵的奇异值，并且计算条件数 σ_1/σ_4.
29. [M]计算 2.3 节习题 10 中 5×5 矩阵的奇异值，并且计算条件数 σ_1/σ_5.

练习题答案

如果 $A=U\Sigma V^{\mathrm{T}}$，此处Σ是 $m\times n$ 矩阵，那么 $A^{\mathrm{T}}=(V^{\mathrm{T}})^{\mathrm{T}}\Sigma^{\mathrm{T}}U^{\mathrm{T}}=V\Sigma^{\mathrm{T}}U^{\mathrm{T}}$，这是 A^{T} 的一个奇异值分解，原因是 V 和 U 是正交矩阵，且 Σ^{T} 是一个 $n\times m$ “对角”矩阵，由于Σ和 Σ^{T} 具有同样的非零对角元素，A 和 A^{T} 具有同样的非零奇异值.（注意：如果 A 是 $2\times n$ 矩阵，那么 AA^{T} 仅为 2×2 矩阵，且它的特征值比 $A^{\mathrm{T}}A$ 的特征值更容易（手工）计算.）

7.5 图像处理和统计学中的应用

本章介绍性实例中的卫星图像问题给出一个多维或多变量数据的例子，组织数据后使得数据集合中的每组数据可看成是 $\mathbb{R}^n$ 的点（向量），本节的主要目标是介绍**主成分分析**的方法，用于分析这类多维数据的技巧. 计算将说明正交对角化和奇异值分解的应用方法.

主成分分析可用于任何数据，包括测量清单上采集的对象或个体. 例如，研究生产塑料材料的化学过程时，为了监控生产过程，在材料生产过程中取得 300 样本，且每一个样本经过 8 个一组的测试. 如熔化点、密度、粘性、抗拉强度等，实验室的每一个样本报告是一个属于 $\mathbb{R}^8$ 的向量，且这类向量集合形成一个 8×300 的矩阵，称为**观测矩阵**.

粗略地讲，我们可以说控制过程的数据是 8 维的，下面两个例子中描述的数据可以用图形给出.

例 1 一个二维数据的例子是，N 个大学生关于体重和身高的一组数据，令 X_j 表示 $\mathbb{R}^2$ 中的**观测向量**，它列出第 j 个学生的体重和身高，如果用 w 表示体重，h 表示身高，那么观测矩阵的形式为：

$$\begin{bmatrix} w_1 & w_2 & \cdots & w_N \\ h_1 & h_2 & \cdots & h_N \end{bmatrix}$$
$$\begin{matrix} \uparrow & \uparrow & & \uparrow \\ \boldsymbol{X}_1 & \boldsymbol{X}_2 & & \boldsymbol{X}_N \end{matrix}$$

观测向量的集合可以形象地表示为一个二维散列图，见图 7-18. ■

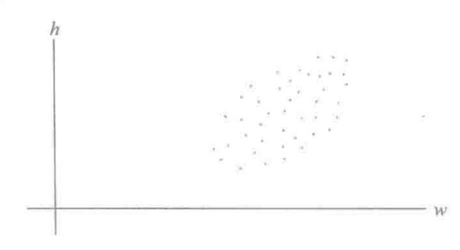


图 7-18 观测向量 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_N$ 的散列图

例 2 本章介绍性实例中，关于内华达州铁路峡谷的前三幅图可以作为某区域的具有三个光谱分量的图形，原因是它们在三个独立的波长同时给出该区域的度量，每幅图给出同一自然区域的不同信息. 例如，每幅图中位于左上角的第一像素对应地面的同一位置（大约 30 米×30 米）. 每一个像素对应着 $\mathbb{R}^3$ 的一个观测向量，它列出了该像素在三个光谱段中的信号强度.

典型的图形是 2000×2000 像素，使得图像有 400 万像素，图像的数据形成一个 3 行和 400 万列的矩阵（列可以调整为方便的次序）. 在这种情形下，数据的“多维”特征是指 3 个光谱维数，而不是自然属于图形的二维空间维数. 数据也许如图 7-19 所示，形象地表示为 $\mathbb{R}^3$ 中的 400 万个点.

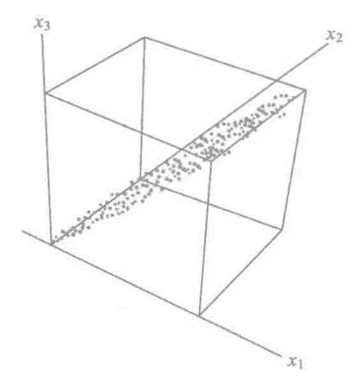


图 7-19 一幅卫星图像中光谱数据的散列图 ■

均值和协方差

为准备主成分分析，令$[\boldsymbol{X}_1 \ \cdots \ \boldsymbol{X}_N]$是如上描述的一个 $p\times N$ 观测矩阵，观测向量 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_N$ 的**样本均值 $\boldsymbol{M}$** 由下式给出：

$$\boldsymbol{M}=\frac{1}{N}(\boldsymbol{X}_1+\cdots+\boldsymbol{X}_N)$$

对图 7-18 中的数据，样本均值是散列图的中心，对 $k=1,\cdots,N$ ，令

$$\hat{\boldsymbol{X}}_k=\boldsymbol{X}_k-\boldsymbol{M}$$

$p\times N$ 矩阵的列

$$B=[\hat{\boldsymbol{X}}_1\ \hat{\boldsymbol{X}}_2\cdots\ \hat{\boldsymbol{X}}_N]$$

具有零样本均值，这样的 B 称为**平均偏差形式**. 当图 7-18 中的数据减去样本均值后，得到的散列图具有图 7-20 的形式.

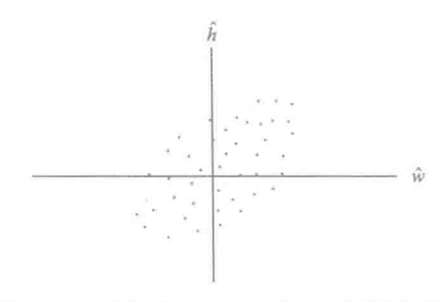


图 7-20　平均偏差形式的体重-身高数据

（样本）**协方差矩阵**是一个 $p\times p$ 矩阵 S ，其定义为

$$S=\frac{1}{N-1}BB^{\mathrm{T}}$$

由于任何具有 BB^{T} 形式的矩阵是半正定的，所以 S 也是半正定的.（见 7.2 节的习题 25，互换 B 和 B^{T}.）

例 3　从一个总体中随机取出 4 个样本作三次测量，每一个样本的观测向量为：

$$\boldsymbol{X}_1=\begin{bmatrix}1\\2\\1\end{bmatrix},\ \boldsymbol{X}_2=\begin{bmatrix}4\\2\\13\end{bmatrix},\ \boldsymbol{X}_3=\begin{bmatrix}7\\8\\1\end{bmatrix},\ \boldsymbol{X}_4=\begin{bmatrix}8\\4\\5\end{bmatrix}$$

计算样本均值和协方差矩阵.

解　样本均值是

$$\boldsymbol{M}=\frac{1}{4}\left(\begin{bmatrix}1\\2\\1\end{bmatrix}+\begin{bmatrix}4\\2\\13\end{bmatrix}+\begin{bmatrix}7\\8\\1\end{bmatrix}+\begin{bmatrix}8\\4\\5\end{bmatrix}\right)=\frac{1}{4}\begin{bmatrix}20\\16\\20\end{bmatrix}=\begin{bmatrix}5\\4\\5\end{bmatrix}$$

从 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_4$ 中减去样本均值，我们得到：

$$\hat{X}_1=\begin{bmatrix}-4\\-2\\-4\end{bmatrix},\ \hat{X}_2=\begin{bmatrix}-1\\-2\\8\end{bmatrix},\ \hat{X}_3=\begin{bmatrix}2\\4\\-4\end{bmatrix},\ \hat{X}_4=\begin{bmatrix}3\\0\\0\end{bmatrix}$$

并且

$$B=\begin{bmatrix}-4&-1&2&3\\-2&-2&4&0\\-4&8&-4&0\end{bmatrix}$$

样本协方差矩阵为

$$S=\frac{1}{3}\begin{bmatrix}-4&-1&2&3\\-2&-2&4&0\\-4&8&-4&0\end{bmatrix}\begin{bmatrix}-4&-2&-4\\-1&-2&8\\2&4&-4\\3&0&0\end{bmatrix}$$

$$=\frac{1}{3}\begin{bmatrix}30&18&0\\18&24&-24\\0&-24&96\end{bmatrix}=\begin{bmatrix}10&6&0\\6&8&-8\\0&-8&32\end{bmatrix}$$

■

为了讨论 $S=[s_{ij}]$ 中的元素，令 $\boldsymbol{X}$ 表示在观测向量集合中变化的向量，且 $x_1,\cdots,x_p$ 表示 $\boldsymbol{X}$ 的坐标，那么例如，x_1 是一个在 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_N$ 集合中变化的第一个坐标的数值，对 $j=1,\cdots,p$，S 中的对角元素 s_{jj} 称为 x_j 的**方差**.

x_j 的方差用来度量 x_j 值的分散性（见习题 13），在例 3 中，x_1 的方差是 10，x_3 的方差是 32，32 大于 10 的事实说明，对应向量第三个元素的集合包含比第一个元素的集合更大的取值范围.

数据的**总方差**是指 S 中对角线上方差的总和. 一般地，一个方阵 S 中对角元素之和称为矩阵的**迹**，记作 $\mathrm{tr}(S)$，这样

$$\{总方差\}=\mathrm{tr}(S)$$

S 中的元素 s_{ij}（$i\neq j$）称为 x_i 和 x_j 的**协方差**. 观察例 3 中的协方差，x_1 和 x_3 的协方差是零，因为 S 中的(1,3)元素是零，统计学家称 x_1 和 x_3 是**无关**的. 如果大部分或所有变量 $x_1,\cdots,x_p$ 是无关的，即当 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_N$ 的协方差矩阵是对角阵或几乎是对角阵时，则 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_N$ 中多变量数据的分析可以简化.

主成分分析

为简单起见，假设矩阵 $[\boldsymbol{X}_1\ \cdots\ \boldsymbol{X}_N]$ 已经是平均偏差形式，主成分分析的目标是找到一个 $p\times p$ 正交矩阵 $P=[\boldsymbol{u}_1\ \cdots\ \boldsymbol{u}_p]$，确定一个变量代换 $\boldsymbol{X}=P\boldsymbol{Y}$，或

$$\begin{bmatrix}x_1\\x_2\\\vdots\\x_p\end{bmatrix}=[\boldsymbol{u}_1\ \boldsymbol{u}_2\cdots\boldsymbol{u}_p]\begin{bmatrix}y_1\\y_2\\\vdots\\y_p\end{bmatrix}$$

并具有新的变量 $y_1,\cdots,y_p$ 两两无关的性质，且整理后的方差具有递减顺序.

变量的正交变换 $\boldsymbol{X}=P\boldsymbol{Y}$ 说明，每一个观测向量 $\boldsymbol{X}_k$ 得到一个“新名称” $\boldsymbol{Y}_k$，使得 $\boldsymbol{X}_k=P\boldsymbol{Y}_k$，注意到 $\boldsymbol{Y}_k$ 是 $\boldsymbol{X}_k$ 相对于以 P 的列为基的坐标向量，且对 $k=1,\cdots,N$ 有 $\boldsymbol{Y}_k=P^{-1}\boldsymbol{X}_k=P^{\mathrm{T}}\boldsymbol{X}_k$.

不难验证，对任何正交矩阵 P，$\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_N$ 的协方差是 $P^{\mathrm{T}}SP$（习题 11）. 于是，期望的正交矩阵 P 是一矩阵使得 $P^{\mathrm{T}}SP$ 为对角形. 设 D 是对角形矩阵且 S 的特征值 $\lambda_1,\cdots,\lambda_p$ 位于对角线上，重新整理使得 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_p\geqslant 0$，并记 P 是正交矩阵，它的列是对应单位特征向量 $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p$，那么 $S=PDP^{\mathrm{T}}$ 且有 $P^{\mathrm{T}}SP=D$.

协方差矩阵 S 的单位特征向量 $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p$ 称为（观测矩阵中的）数据的**主成分**. **第一主成分**是 S 中最大特征值对应的特征向量. **第二主成分**是 S 中第二大特征值对应的特征向量，以此类推.

第一主成分 $\boldsymbol{u}_1$ 可用下列方式确定新变量 y_1，设 $c_1,\cdots,c_p$ 是 $\boldsymbol{u}_1$ 中的元素，由于 $\boldsymbol{u}_1^{\mathrm{T}}$ 是 P^{T} 的行，方程 $\boldsymbol{Y}=P^{\mathrm{T}}\boldsymbol{X}$ 表明

$$y_1=\boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{X}=c_1x_1+c_2x_2+\cdots+c_px_p$$

于是 y_1 是原变量 $x_1,\cdots,x_p$ 的线性组合，并用特征向量 $\boldsymbol{u}_1$ 中的元素作为权值. 同样的方式，$\boldsymbol{u}_2$ 确定变量 y_2，以此类推.

例 4 铁路峡谷（例 2）的多谱图像的初始数据包含 $\mathbb{R}^3$ 中 400 万个向量，其协方差矩阵是[㊀]

$$S=\begin{bmatrix}2\,382.78 & 2\,611.84 & 2\,136.20\\ 2\,611.84 & 3\,106.47 & 2\,553.90\\ 2\,136.20 & 2\,553.90 & 2\,650.71\end{bmatrix}$$

求数据的主成分，且列出由第一主成分确定的新变量.

解 S 的特征值和相关的主成分（单位特征向量）是

$$\lambda_1=7\,614.23\quad \lambda_2=427.63\quad \lambda_3=98.10$$

$$\boldsymbol{u}_1=\begin{bmatrix}0.541\,7\\0.629\,5\\0.557\,0\end{bmatrix}\quad \boldsymbol{u}_2=\begin{bmatrix}-0.489\,4\\-0.302\,6\\0.817\,9\end{bmatrix}\quad \boldsymbol{u}_3=\begin{bmatrix}0.683\,4\\-0.715\,7\\0.144\,1\end{bmatrix}$$

为简化而用小数点后两位小数，第一主成分的变量是

$$y_1=0.54x_1+0.63x_2+0.56x_3$$

在本章介绍性实例中，这个方程用于生成图 7-1d. 变量 x_1,x_2,x_3 是三个光谱段中的信号强度. 将 x_1 的值转化为介于黑色和白色之间的灰度，生成图 7-1a. 类似地，x_2 和 x_3 的值分别生成图 7-1b 和 c. 对图 7-1d 中的每一个像素，用 y_1 计算得到灰度值，即加权的 x_1,x_2,x_3 的线性组合. 在这个意义下，图 7-1d“显示”数据的第一主成分. ■

在例 4 中，变换后数据的协方差矩阵，用变量 y_1,y_2,y_3 表示为：

$$D=\begin{bmatrix}7\,614.23 & 0 & 0\\ 0 & 427.63 & 0\\ 0 & 0 & 98.10\end{bmatrix}$$

㊀ 例 4 和习题 5、习题 6 中的数据由马里兰州罗克维尔市地球卫星公司提供.

尽管 D 比原来的协方差矩阵 S 明显简单，但构造新变量的优点仍然不明显. 然而，变量 y_1, y_2, y_3 的方差出现在对角矩阵 D 的对角线上，并且明显看出 D 中第一个方差比其余两个大得多. 如我们将要看到的，这个事实允许我们将数据当作一维而不是三维的.

多变量数据的降维

对大多数数据的变化或动态范围，当新变量 $y_1,\cdots,y_p$ 中的一些变量的变化较小时，主成分分析有潜在的应用价值.

可以证明变量的正交变换，$X = PY$，不改变数据的总方差.（粗略地讲，这个结论是真的，其原因是左乘 P 不改变向量的长度或它们的夹角，见习题 12.）这说明，如果 $S = PDP^{\mathrm{T}}$，那么

$$\{x_1,\cdots,x_p\text{的总方差}\} = \{y_1,\cdots,y_p\text{的总方差}\} = \mathrm{tr}(D) = \lambda_1 + \cdots + \lambda_p$$

y_j 的方差是 λ_j，且商 $\lambda_j/\mathrm{tr}(S)$ 度量总体方差成分中，被 y_j “说明”或“记录”的比例.

例 5 计算铁路峡谷例题中各种方差占总方差的成分，显示主成分的多光谱数据，在本章介绍性实例的图 7-1d~f 已经画出.

解 数据的总方差是

$$\mathrm{tr}(D) = 7\,614.23 + 427.63 + 98.10 = 8\,139.96$$

（可验证这个数等于 $\mathrm{tr}(S)$.）主成分在总方差的百分比分别是

$$\underset{\text{第一成分}}{\frac{7\,614.23}{8\,139.96} = 93.5\%} \qquad \underset{\text{第二成分}}{\frac{427.63}{8\,139.96} = 5.3\%} \qquad \underset{\text{第三成分}}{\frac{98.10}{8\,139.96} = 1.2\%}$$

其意义是，地球资源卫星关于铁路峡谷地区收集的 93.5%信息显示在图 7-1d 上，5.3%显示在图 7-1e 上，而仅有剩余的 1.2%显示在图 7-1f 中. ■

例 5 的计算表明，数据的第三个坐标实际上没有变化，y_3 的值几乎接近于零，从几何意义上看，数据点几乎位于平面 $y_3 = 0$，且它们的位置可由已知的 y_1，y_2 相当精确地确定，实际上，y_2 的方差也相对很小，这说明点集几乎位于一条直线上，数据几乎是一维的，见图 7-19，其数据像一个冰棒棍.

主成分变量的特征

如果 $y_1,\cdots,y_p$ 是来自一个 $P\times N$ 观测矩阵的主成分分析，那么 y_1 的方差在下列意义下可能尽量大，如果 $\boldsymbol{u}$ 是任意一个单位向量且 $y = \boldsymbol{u}^{\mathrm{T}}\boldsymbol{X}$，那么当 X 在原来数据 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_N$ 变化范围时，y 的方差值为 $\boldsymbol{u}^{\mathrm{T}}S\boldsymbol{u}$，由 7.3 节的定理 8，在所有单位向量 $\boldsymbol{u}$ 中，$\boldsymbol{u}^{\mathrm{T}}S\boldsymbol{u}$ 的最大值就是 S 的最大特征值 λ_1 且这个方差可以在 $\boldsymbol{u}$ 等于对应的特征向量 $\boldsymbol{u}_1$ 处达到. 同样的方式，定理 8 表明 y_2 的方差最大值，可能出现在与 y_1 无关的所有变量 $y = \boldsymbol{u}^{\mathrm{T}}\boldsymbol{X}$ 中，同样，y_3 的方差最大值可能出现在与 y_1 和 y_2 都无关的所有变量中，以此类推.

数值计算的注解 在实际应用中，奇异值分解是进行主成分分析的主要工具，如果 B 是一个具有平均偏差形式的 $p\times N$ 观测矩阵，且如果 $A = (1/\sqrt{N-1})B^{\mathrm{T}}$，那么 $A^{\mathrm{T}}A$ 是协方差矩阵 S. A 的奇异值的平方是 S 的 p 个特征值，且 A 的右奇异向量是数据的主成分.

像 7.4 节所提到的，A 的奇异值分解迭代计算比 S 的特征值分解更快更准确. 在例如本章介绍性实例中提到的超谱图像处理（$p = 224$）中，更是正确. 在专业化的工

作站上，主成分分析可在数秒内完成.

进一步阅读材料

Lillesand, Thomas M., and Ralph W. Kiefer, *Remote Sensing and Image Interpretation*, 4th ed. (New York: John Wiley, 2000).

练习题

下表列出 5 个男孩子的体重和身高.

男　孩	#1	#2	#3	#4	#5
体重（lb[①]）	120	125	125	135	145
身高（in[②]）	61	60	64	68	72

① 1 lb=0.453 592 37 kg.——编辑注

② 1 in=0.025 4 m.——编辑注

1. 求数据的协方差矩阵.
2. 给出数据的主成分分析，求一个数量指标来解释大多数数据中的变化.

习题 7.5

对习题 1 和习题 2，将观测矩阵变为平均偏差形式，并构造样本的协方差矩阵.

1. $\begin{bmatrix} 19 & 22 & 6 & 3 & 2 & 20 \\ 12 & 6 & 9 & 15 & 13 & 5 \end{bmatrix}$

2. $\begin{bmatrix} 1 & 5 & 2 & 6 & 7 & 3 \\ 3 & 11 & 6 & 8 & 15 & 11 \end{bmatrix}$

3. 求习题 1 中数据的主成分.

4. 求习题 2 中数据的主成分.

5. [M]一个地球资源卫星有三个光谱分量，由佛罗里达州 Homestead 空军基地（自从 1992 年该基地被安德鲁飓风袭击后），所构造数据的协方差矩阵显示在下面，求数据的第一主成分分量，并且计算这个成分在总体方差中的百分比.
$$S=\begin{bmatrix} 164.12 & 32.73 & 81.04 \\ 32.73 & 539.44 & 249.13 \\ 81.04 & 249.13 & 189.11 \end{bmatrix}$$

6. [M]下面的协方差矩阵来自华盛顿哥伦比亚河地球资源卫星的图像，采用数据的三个光谱段，令 x_1，x_2，x_3 表示图像中的每个像素的光谱成分，求新的变量形式 $y_1=c_1x_1+c_2x_2+c_3x_3$ 具有最大可能的方差，限制条件是 $c_1^2+c_2^2+c_3^2=1$，y_1 中的数据占总方差的百分比是多少？
$$S=\begin{bmatrix} 29.64 & 18.38 & 5.00 \\ 18.38 & 20.82 & 14.06 \\ 5.00 & 14.06 & 29.21 \end{bmatrix}$$

7. 令 x_1，x_2 表示习题 1 中二维数据变量，求一个形如 $y_1=c_1x_1+c_2x_2$ 的新变量 y_1，满足条件 $c_1^2+c_2^2=1$，在给定数数据下，使得 y_1 的方差尽可能具有最大值，y_1 中的数据占总方差的百分比是多少？

8. 对习题 2 的数据重复习题 7 的问题.

9. 假设三个实验中，大学生的随机样本已经掌握，设 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_N$ 是 $\mathbb{R}^3$ 中的观测向量，且对 $j=1$，2，3, x_j 表示第 j 次考试中学生的分数，设数据的协方差矩阵是
$$S=\begin{bmatrix} 5 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 7 \end{bmatrix}$$
设 y 表示学生表现的指标，且 $y_1=c_1x_1+c_2x_2+c_3x_3$ 和 $c_1^2+c_2^2+c_3^3=1$，选取 c_1，c_2，c_3，使得 y 在数据集变化时，y 的方差尽可能大（提示：协方差矩阵的特征值是 $\lambda=3$, 6 和 9）.

10. [M]对 $S=\begin{bmatrix}5 & 4 & 2\\ 4 & 11 & 4\\ 2 & 4 & 5\end{bmatrix}$，重复习题 9 的问题.

11. 给定平均偏差形式的多变量数据 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_N$（属于 $\mathbb{R}^p$），P 是 $p \times p$ 矩阵，且对 $k=1$，$2, \cdots, N$，定义 $\boldsymbol{Y}_k = P^{\mathrm{T}}\boldsymbol{X}_k$.
    a. 证明：$\boldsymbol{Y}_1$, $\boldsymbol{Y}_2$, $\cdots, \boldsymbol{Y}_N$ 是平均偏差形式.（提示：令 $\boldsymbol{w}$ 是 $\mathbb{R}^N$ 中的向量且每一个元素为 1，那么 $[\boldsymbol{X}_1 \cdots \boldsymbol{X}_N]\boldsymbol{w} = \boldsymbol{0}$（$\mathbb{R}^p$ 中的零向量）.）
    b. 证明：如果 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_N$ 的协方差矩阵是 S，那么 $\boldsymbol{Y}_1$, $\boldsymbol{Y}_2$, $\cdots, \boldsymbol{Y}_N$ 的协方差矩阵是 $P^{\mathrm{T}}SP$.

12. 令 $\boldsymbol{X}$ 表示一个 $p \times N$ 观测矩阵的一个列向量，且令 P 是一个 $p \times p$ 正交矩阵，证明：变量变换 $\boldsymbol{X}=P\boldsymbol{Y}$ 不影响数据的总方差.（提示：由习题 11，只需证明 $\mathrm{tr}(P^{\mathrm{T}}SP)=\mathrm{tr}(S)$，利用 5.4 节习题 25 提到的轨迹的性质.）

13. 样本协方差矩阵是一个样本的 N 个测量数值的方差公式的一般形式，如 $t_1, \cdots, t_N$，如果 $\boldsymbol{m}$ 是 $t_1, \cdots, t_N$ 的平均值，那么样本方差由下式给出：

$$\frac{1}{N-1}\sum_{k=1}^{n}(t_k - m)^2 \qquad (1)$$

证明前面例 3 中定义的样本协方差矩阵 S，可以写成类似（1）的形式.（提示：利用分块矩阵的乘法，将 S 写成 $1/(N-1)$ 乘 N 个 $p \times p$ 矩阵的和，对 $1 \leqslant k \leqslant N$，用 $\boldsymbol{X}_k - \boldsymbol{M}$ 代替 $\hat{\boldsymbol{X}}_k$.）

练习题答案

1. 首先将数据整理为平均偏差形式，容易看出样本的平均向量是 $\boldsymbol{M}=\begin{bmatrix}130\\65\end{bmatrix}$，从观测向量（表中的列）中减去 $\boldsymbol{M}$ 可以得到

$$B=\begin{bmatrix}-10 & -5 & -5 & 5 & 15\\ -4 & -5 & -1 & 3 & 7\end{bmatrix}$$

那么样本协方差矩阵是

$$S=\frac{1}{5-1}\begin{bmatrix}-10 & -5 & -5 & 5 & 15\\ -4 & -5 & -1 & 3 & 7\end{bmatrix}\begin{bmatrix}-10 & -4\\ -5 & -5\\ -5 & -1\\ 5 & 3\\ 15 & 7\end{bmatrix}=\frac{1}{4}\begin{bmatrix}400 & 190\\ 190 & 100\end{bmatrix}=\begin{bmatrix}100.0 & 47.5\\ 47.5 & 25.0\end{bmatrix}$$

2. S 的特征值是（精确到两位小数）

$$\lambda_1 = 123.02 \text{ 和 } \lambda_2 = 1.98$$

对应 λ_1 的单位特征向量是 $\boldsymbol{u}=\begin{bmatrix}0.900\\0.436\end{bmatrix}$.（由于 S 是 2×2 矩阵，如果没有矩阵程序，则可以用手算.） 为计算数量指标，取

$$y = 0.900\hat{w} + 0.436\hat{h}$$

此处 $\hat{w}$ 和 $\hat{h}$ 分别表示平均偏差形式的体重和身高. 在数据集合中该指标的方差是 123.02. 由于总方差 $\mathrm{tr}(S)=100+25=125$，数量指标解释了几乎所有（98.4%）的数据方差.

练习题 1 的原始数据和由第一主成分 $\boldsymbol{u}$ 确定的直线，都显示在图 7-21 中（用参数向量形式，直线是 $\boldsymbol{x}=\boldsymbol{M}+t\boldsymbol{u}$），在到直线的垂直距离的平方和最小的意义下，可以证明直线是数据的最佳逼近. 事实上，主成分分析与术语*正交回归*是等同的，但正交回归完属于另外的内容，以后我们会遇到.

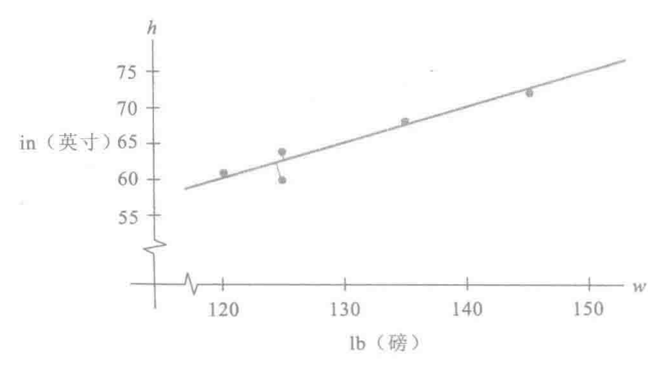


图 7-21　一个由数据的第一主成分确定的正交回归直线

补充习题

1. 说明每一个论断是正确还是错误的，验证你的答案，在每一部分中，A 表示一个 $n\times n$ 矩阵.
   a. 如果 A 是可正交对角化矩阵，那么 A 是对称的.
   b. 如果 A 是一个正交矩阵，那么 A 是对称的.
   c. 如果 A 一个正交矩阵，那么 $\|Ax\|=\|x\|$，对所有 $\mathbb{R}^n$ 的向量均成立.
   d. 二次型 $x^{\mathrm{T}}Ax$ 的主轴可以是任意将 A 对角化的矩阵 P 的列.
   e. 如果 P 是一个 $n\times n$ 具有正交列的矩阵，那么 $P^{\mathrm{T}}=P^{-1}$.
   f. 如果二次型的每一个系数是正的，那么二次型是正定的.
   g. 如果对某些 x，有 $x^{\mathrm{T}}Ax>0$，那么二次型 $x^{\mathrm{T}}Ax$ 是正定的.
   h. 通过适当的变量变换，任何二次型可以变成一个没有交叉乘积项的二次型.
   i. 二次型 $x^{\mathrm{T}}Ax$ 在 $\|x\|=1$ 条件下的最大值是矩阵 A 对角线上的最大元素.
   j. 一个正定二次型 $x^{\mathrm{T}}Ax$ 的最大值是 A 的最大特征值.
   k. 一个正定二次型可以通过一个合适的变量变换 $x=Pu$ 变为负定二次型，其中 P 是正交矩阵.
   l. 一个不定二次型是指特征值不定的二次型.
   m. 如果 P 是 $n\times n$ 正交矩阵，那么变量变换 $x=Pu$ 将 $x^{\mathrm{T}}Ax$ 变换为矩阵为 $P^{-1}AP$ 的二次型.
   n. 如果 U 是 $m\times n$ 且列正交的矩阵，那么 $UU^{\mathrm{T}}x$ 是 x 在 $\operatorname{Col}U$ 上的正交投影.
   o. 如果 B 是 $m\times n$ 矩阵且 x 是 $\mathbb{R}^n$ 中的一个单位向量，那么 $\|Bx\|\leqslant\sigma_1$，此处 σ_1 是 B 的第一个奇异值.
   p. 一个 $m\times n$ 矩阵 B 的奇异值分解，可以写成 $B=P\Sigma Q$，此处，P 是一个 $m\times m$ 正交矩阵，Q 是一个 $n\times n$ 正交矩阵，Σ 是一个 $m\times n$ “对角”矩阵.
   q. 如果 A 是 $n\times n$ 矩阵，那么 A 和 $A^{\mathrm{T}}A$ 具有同样的奇异值.
2. 令 $\{u_1,\cdots,u_n\}$ 是 $\mathbb{R}^n$ 的标准正交基，且 $\lambda_1,\cdots,\lambda_n$ 是任何实数，定义
$$A=\lambda_1u_1u_1^{\mathrm{T}}+\cdots+\lambda_nu_nu_n^{\mathrm{T}}$$
   a．证明 A 是对称的.
   b．证明 $\lambda_1,\cdots,\lambda_n$ 是 A 的特征值.
3. 令 A 是秩为 r 的 $n\times n$ 对称矩阵，试解释，为什么 A 的谱分解是将 A 表示为 r 个秩为 1 的矩阵之和.

4. 令 A 是一个 $n\times n$ 对称矩阵.
   a. 证明 $(\mathrm{Col}\ A)^{\perp}=\mathrm{Nul}\ A$.（提示：见 6.1 节.）
   b. 证明：对每一个属于 $\mathbb{R}^n$ 的 $\boldsymbol{y}$，可以写成 $\boldsymbol{y}=\hat{\boldsymbol{y}}+\boldsymbol{z}$ 的形式，其中 $\hat{\boldsymbol{y}}$ 属于 Col A，$\boldsymbol{z}$ 属于 Nul A.
5. 证明：如果 $\boldsymbol{v}$ 是 $n\times n$ 矩阵 A 的一个特征向量，且 λ 是对应的 A 的非零特征值，那么 $\boldsymbol{v}$ 属于 Col A.（提示：利用特征向量的定义.）
6. 设 A 是一个 $n\times n$ 对称矩阵，利用习题 5 和 $\mathbb{R}^n$ 的一个特征向量基，给出习题 4 (b) 中分解的第二种证明.
7. 证明：一个 $n\times n$ 矩阵 A 是正定的充分必要条件是，A 有一个楚列斯基分解，即 $A=R^{\mathrm{T}}R$ 对一些可逆上三角形矩阵 R 成立，其中 R 的主对角元素都是正数.（提示：利用 QR 分解和 7.2 节习题 26）.
8. 利用习题 7 证明：如果 A 是正定的，那么 A 具有一个 LU 分解 $A=LU$，此处 U 的对角线上有正主元.（反之也真.）

如果 A 是 $m\times n$ 矩阵，那么矩阵 $G=A^{\mathrm{T}}A$ 称为 A 的格拉姆矩阵，在这种情形，G 的元素是 A 的列向量的内积.

9. 证明：任意矩阵 A 的格拉姆矩阵是半正定的，且与 A 有相同的秩.（见 6.5 节的习题.）
10. 证明：如果一个 $n\times n$ 矩阵 G 是半正定且有秩 r，那么 G 是某 $r\times n$ 矩阵 A 的格拉姆矩阵，称为 G 的显秩矩阵的分解.（提示：考虑 G 的谱分解，首先将 G 写成一个 $n\times r$ 矩阵 B 的 BB^{T} 形式.）
11. 证明：任何 $n\times n$ 矩阵 A 有一个形如 $A=PQ$ 的极分解，此处 P 是 $n\times n$ 半正定矩阵且与 A 具有相同的秩，Q 是一个 $n\times n$ 正交矩阵.（提示：利用奇异值分解，$A=U\Sigma V^{\mathrm{T}}$，且观察到 $A=(U\Sigma U^{\mathrm{T}})(UV^{\mathrm{T}})$.）这种分解用于例如机械工程中给材料的变形建模. 矩阵 P 表示材料（沿 P 的特征向量的方向上）的拉伸或压缩变换，Q 表示材料在空间的旋转变换.

习题 12~14 涉及一个 $m\times n$ 矩阵 A，且具有简化的奇异值分解，$A=U_rDV_r^{\mathrm{T}}$，且其伪逆是 $A^{+}=V_rD^{-1}U_r^{\mathrm{T}}$.

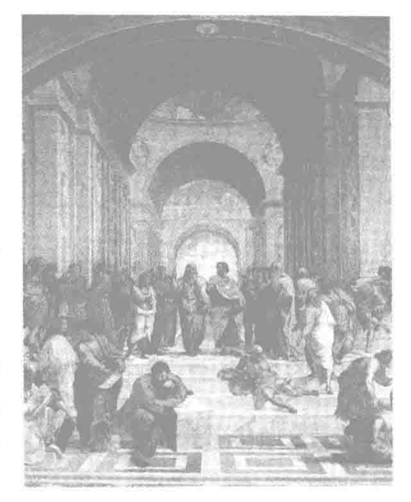
12. 验证 A^{+} 的性质：
    a. 对每个属于 $\mathbb{R}^m$ 的 $\boldsymbol{y}$，$AA^{+}\boldsymbol{y}$ 是 $\boldsymbol{y}$ 在 Col A 上的正交投影.
    b. 对每个属于 $\mathbb{R}^n$ 的 $\boldsymbol{x}$，$A^{+}A\boldsymbol{x}$ 是 $\boldsymbol{x}$ 在 Row A 上的正交投影.
    c. $AA^{+}A=A$ 和 $A^{+}AA^{+}=A^{+}$.
13. 若方程 $A\boldsymbol{x}=\boldsymbol{b}$ 是相容的，且令 $\boldsymbol{x}^{+}=A^{+}\boldsymbol{b}$. 由 6.3 节的习题 23，存在唯一一个属于 Row A 的向量 $\boldsymbol{p}$，使得 $A\boldsymbol{p}=\boldsymbol{b}$，下面的步骤证明 $\boldsymbol{x}^{+}=\boldsymbol{p}$ 且 $\boldsymbol{x}^{+}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小长度解.
    a. 证明：$\boldsymbol{x}^{+}$ 属于 Row A.（提示：将 $\boldsymbol{b}$ 写成关于 $\boldsymbol{x}$ 的 $A\boldsymbol{x}$ 形式，并且利用习题 12.）
    b. 证明：$\boldsymbol{x}^{+}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的一个解.
    c. 证明：如果 $\boldsymbol{u}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的任意一个解，那么 $\|\boldsymbol{x}^{+}\|\leqslant\|\boldsymbol{u}\|$. 且只有当 $\boldsymbol{u}=\boldsymbol{x}^{+}$ 时等号成立.
14. 给定 $\mathbb{R}^m$ 中的任何 $\boldsymbol{b}$，修改习题 13 证明：$A^{+}\boldsymbol{b}$ 是最小长度的最小二乘解.（提示：考虑方程 $A\boldsymbol{x}=\hat{\boldsymbol{b}}$，此处 $\hat{\boldsymbol{b}}$ 是 $\boldsymbol{b}$ 在 Col A 的正交投影.）

[M]对习题 15 和习题 16 构造 A 的伪逆. 开始时使用矩阵程序计算的奇异值分解，或者，如果没有这个程序，首先对 $A^{\mathrm{T}}A$ 进行正交对角化. 使用伪逆求解 $A\boldsymbol{x}=\boldsymbol{b}$，其中 $\boldsymbol{b}=(6,-1,-4,6)$，并记 $\hat{\boldsymbol{x}}$ 是所求的解，用计算方法验证 $\hat{\boldsymbol{x}}$ 属于 Row A，求一个属于 Nul A 的非零向量 $\boldsymbol{u}$，并验证 $\|\hat{\boldsymbol{x}}\|<\|\hat{\boldsymbol{x}}+\boldsymbol{u}\|$，由习题 13（c）可知该式一定成立.

15. $A=\begin{bmatrix}-3 & -3 & -6 & 6 & 1\\ -1 & -1 & -1 & 1 & -2\\ 0 & 0 & -1 & 1 & -1\\ 0 & 0 & -1 & 1 & -1\end{bmatrix}$
16. $A=\begin{bmatrix}4 & 0 & -1 & -2 & 0\\ -5 & 0 & 3 & 5 & 0\\ 2 & 0 & -1 & -2 & 0\\ 6 & 0 & -3 & -6 & 0\end{bmatrix}$

第 8 章　向量空间的几何学

介绍性实例　柏拉图多面体



在公元前 387 年的雅典城，希腊哲学家柏拉图创建了世界上第一所大学. 课程涉及天文学、生物学、政治学理论和哲学. 这些学科中最受柏拉图重视的是几何学，这一事实可以从学院门上写的一句话得到验证，这句话是这样写的：“不要让不懂几何的人进入我的门.”

希腊人被几何形状深深地吸引了，如正多面体. 如果一个多面体的面都是全等的正多边形，并且所有顶点处的角都是相等的，那么称这样的多面体为正多面体. 在早于欧几里得的 150 年前，毕达哥拉斯知道至少三种正多面体：四面体（4 个三角面）、立方体（6 个方形面）和八面体（8 个三角面）（见图 8-1）. 只有 5 种这样的多面体自然地出现在常见的矿物晶体中. 另外两种多面体是十二面体(12 个五角面）和二十面体（20 个三角面）.

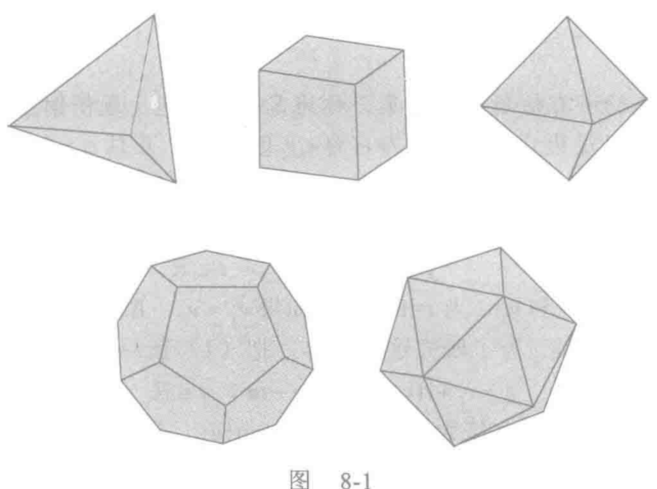


图　8-1

柏拉图在他的著作《Elements》之 XIII 中讨论了这五种多面体的基础理论，从此，这些多面体就带有了他的名字：柏拉图多面体.

从那以后的几个世纪，都无须想象超过三维的几何对象. 但如今，数学家经常处理四维、五维，甚至数百维向量空间中的对象. 而在高维空间中，我们对这些对象的几何性质却未必清楚.

例如，直线在二维空间中所具有的性质和平面在三维空间中所具有的性质，何种性质会在更高维的空间中有用？如何去刻画这些对象？8.1 节和 8.4 节给出了一些回答. 8.4 节中的超平面对理解线性规划问题在高维空间中的特性是十分重要的.

在高于三维的空间中，多面体看起来是什么样的？部分答案可以通过对四维对象做二维投影得到，与三维对象在二维中的投影相似. 8.5 节给出了四维立方体和四维单形的概念.

高维几何理论的学习不仅提供了观察抽象的代数概念的新方法，而且提供了可应用于 $\mathbb{R}^3$ 的新工具. 比如，8.2 节和 8.6 节包含了在计算机画图方面的应用，8.5 节叙述了（习题 21）$\mathbb{R}^3$ 中只存在 5 种正多面体的证明思路.

前几章中大多数的应用涉及子空间的代数计算和向量的线性组合. 这一章研究的是那些被视为几何体的向量集合，如直线段、多边形和实体对象. 独立的向量被看作点. 这里介绍的概念在计算机图形学和其他数学领域都很有用㊀.

纵观整个章节，向量的集合被描述为线性组合，但在组合权重上给了一些限制. 比如，在 8.1 节，权重的和是 1，而在 8.2 节，权重都是正的，且和是 1. 概念的可视化当然是在 $\mathbb{R}^2$ 或 $\mathbb{R}^3$ 中可以实现，但这些概念也应用到 $\mathbb{R}^n$ 和其他向量空间中.

8.1 仿射组合

一个向量的仿射组合就是线性组合的一种特殊形式. 给定 $\mathbb{R}^n$ 中向量（或"点"）$\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p$ 和标量 $c_1,\cdots,c_p$，$\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p$ 的一个仿射组合就是线性组合

$$c_1\boldsymbol{v}_1+\cdots+c_p\boldsymbol{v}_p$$

其权重满足 $c_1+\cdots+c_p=1$.

定义 集合 S 中点的所有仿射组合的集合称为 S 的**仿射包**（或**仿射生成集**），记为 $\operatorname{aff} S$.

一个单点 $\boldsymbol{v}_1$ 的仿射包是集合 $\{\boldsymbol{v}_1\}$，因为它有 $c_1\boldsymbol{v}_1$ 的形式，并且 $c_1=1$. 两个不同点的仿射包通常被写成一种特殊的形式. 假设 $\boldsymbol{y}=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2$，$c_1+c_2=1$. 用 t 代替 c_2，从而 $c_1=1-c_2=1-t$. 则 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 的仿射包就是集合

$$\boldsymbol{y}=(1-t)\boldsymbol{v}_1+t\boldsymbol{v}_2,\quad t\in\mathbb{R} \tag{1}$$

这个点集包括 $\boldsymbol{v}_1$（当 $t=0$ 时）和 $\boldsymbol{v}_2$（当 $t=1$ 时）. 如果 $\boldsymbol{v}_2=\boldsymbol{v}_1$，那么（1）描述了一个点；否则，（1）描述了穿过 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的直线. 为了研究这一内容，把（1）重写为

$$\boldsymbol{y}=\boldsymbol{v}_1+t(\boldsymbol{v}_2-\boldsymbol{v}_1)=\boldsymbol{p}+t\boldsymbol{u},\quad t\in\mathbb{R}$$

这里 $\boldsymbol{p}$ 是 $\boldsymbol{v}_1$，$\boldsymbol{u}$ 是 $\boldsymbol{v}_2-\boldsymbol{v}_1$. $\boldsymbol{u}$ 的所有倍数的集合就是 $\operatorname{Span}\{\boldsymbol{u}\}$，是穿过 $\boldsymbol{u}$ 和原点的直线. 把 $\boldsymbol{p}$ 加到这条直线上的每个点，相当于把穿过 $\boldsymbol{u}$ 和原点的直线平移得到一条穿过 $\boldsymbol{p}$，并与穿过 $\boldsymbol{u}$ 和原点的直线平行的直线，见图 8-2.（比较这个图与 1.5 节中的图 1-25.）

图 8-3 用了原始点 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$，并显示了 $\operatorname{aff}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是穿过 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的直线.

㊀ 见 Foley，van Dam,Feiner,and Hughes，*Computer Graphics-Principles and Practice*, 2nd edition(Boston:Addision-Wesley,1996),pp.1083-1112.其中还讨论了坐标无关的"仿射空间".

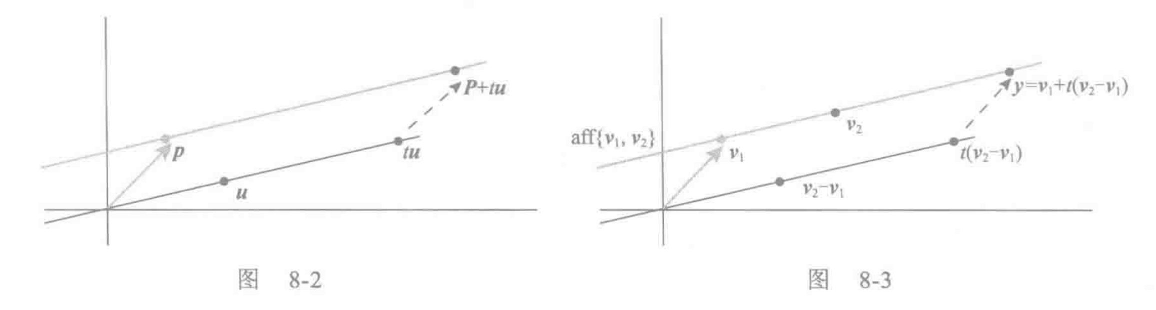


图 8-2　　　　图 8-3

注意，虽然图 8-3 中点 $\boldsymbol{y}$ 是 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的一个仿射组合，但点 $\boldsymbol{y}-\boldsymbol{v}_1$ 等于 $t(\boldsymbol{v}_2-\boldsymbol{v}_1)$，它是 $\boldsymbol{v}_2-\boldsymbol{v}_1$ 的一个线性组合（事实上，是倍乘）. $\boldsymbol{y}$ 和 $\boldsymbol{y}-\boldsymbol{v}_1$ 的这种关系为点的任何仿射组合所共有，如下面定理所述.

定理 1　$\mathbb{R}^n$ 中一个点 $\boldsymbol{y}$ 是 $\mathbb{R}^n$ 中 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 的一个仿射组合，当且仅当 $\boldsymbol{y}-\boldsymbol{v}_1$ 是平移点 $\boldsymbol{v}_2-\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p-\boldsymbol{v}_1$ 的线性组合.

证明　如果 $\boldsymbol{y}-\boldsymbol{v}_1$ 是 $\boldsymbol{v}_2-\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p-\boldsymbol{v}_1$ 的一个线性组合，存在权重 $c_2,\cdots,c_p$ 使得

$$\boldsymbol{y}-\boldsymbol{v}_1=c_2(\boldsymbol{v}_2-\boldsymbol{v}_1)+\cdots+c_p(\boldsymbol{v}_p-\boldsymbol{v}_1) \tag{2}$$

那么

$$\boldsymbol{y}=(1-c_2-\cdots-c_p)\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_p\boldsymbol{v}_p \tag{3}$$

并且这个线性组合中的权重和是 1. 因此，$\boldsymbol{y}$ 是 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 的一个仿射组合. 相反，假定

$$\boldsymbol{y}=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_p\boldsymbol{v}_p \tag{4}$$

这里 $c_1+\cdots+c_p=1$. 由于 $c_1=1-c_2-\cdots-c_p$，方程（4）可写成（3）的形式，并可导出（2），这就证明了 $\boldsymbol{y}-\boldsymbol{v}_1$ 是 $\boldsymbol{v}_2-\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p-\boldsymbol{v}_1$ 的一个线性组合. ■

在定理 1 的叙述中，点 $\boldsymbol{v}_1$ 可以被 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 中任何其他的点所取代. 只要改变证明中的标记就行了.

例 1　令 $\boldsymbol{v}_1=\begin{bmatrix}1\\2\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}2\\5\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}1\\3\end{bmatrix},\boldsymbol{v}_4=\begin{bmatrix}-2\\2\end{bmatrix},\boldsymbol{y}=\begin{bmatrix}4\\1\end{bmatrix}$. 如果可能，把 $\boldsymbol{y}$ 写成 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4$ 的仿射组合.

解　计算平移点

$$\boldsymbol{v}_2-\boldsymbol{v}_1=\begin{bmatrix}1\\3\end{bmatrix},\quad \boldsymbol{v}_3-\boldsymbol{v}_1=\begin{bmatrix}0\\1\end{bmatrix},\quad \boldsymbol{v}_4-\boldsymbol{v}_1=\begin{bmatrix}-3\\0\end{bmatrix},\quad \boldsymbol{y}-\boldsymbol{v}_1=\begin{bmatrix}3\\-1\end{bmatrix}$$

求出标量 c_2,c_3,c_4，使得

$$c_2(\boldsymbol{v}_2-\boldsymbol{v}_1)+c_3(\boldsymbol{v}_3-\boldsymbol{v}_1)+c_4(\boldsymbol{v}_4-\boldsymbol{v}_1)=\boldsymbol{y}-\boldsymbol{v}_1 \tag{5}$$

行化简增广矩阵得：

$$\begin{bmatrix}1&0&-3&3\\3&1&0&-1\end{bmatrix}\sim\begin{bmatrix}1&0&-3&3\\0&1&9&-10\end{bmatrix}$$

这表明（5）是相容的，并且通解是 $c_2=3c_4+3,c_3=-9c_4-10$，c_4 是自由未知量. 当 $c_4=0$ 时，

$$\boldsymbol{y}-\boldsymbol{v}_1=3(\boldsymbol{v}_2-\boldsymbol{v}_1)-10(\boldsymbol{v}_3-\boldsymbol{v}_1)+0(\boldsymbol{v}_4-\boldsymbol{v}_1)$$

即

$$y = 8v_1 + 3v_2 - 10v_3$$

如另选 $c_4 = 1$，则 $c_2 = 6$ 和 $c_3 = -19$，因此，

$$y - v_1 = 6(v_2 - v_1) - 19(v_3 - v_1) + 1(v_4 - v_1)$$

即

$$y = 13v_1 + 6v_2 - 19v_3 + v_4$$

■

虽然例 1 中的步骤对 $\mathbb{R}^n$ 中任意点 $v_1, v_2, \cdots, v_p$ 都适用，但如果选择的点 v_i 是 $\mathbb{R}^n$ 中的一个基就更容易一些了. 例如，令 $\mathcal{B} = \{b_1, \cdots, b_n\}$ 是一个基. 从而 $\mathbb{R}^n$ 中任意 y 是 $b_1, \cdots, b_n$ 的唯一的线性组合. 当且仅当权重和是 1 时，这个线性组合是这些 b_i 的一个仿射组合.（组合的权重正好是 y 在基 $\mathcal{B}$ 下的坐标，如 4.4 节.）

例 2 令 $b_1 = \begin{bmatrix} 4 \\ 0 \\ 3 \end{bmatrix}, b_2 = \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}, b_3 = \begin{bmatrix} 5 \\ 2 \\ 4 \end{bmatrix}, p_1 = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$，$p_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$. 集合 $\mathcal{B} = \{b_1, b_2, b_3\}$ 是 $\mathbb{R}^3$ 的一个基. 判断点 p_1 和 p_2 是否是 $\mathcal{B}$ 的仿射组合中的点.

解 求 p_1 和 p_2 在 $\mathcal{B}$ 下的坐标. 这两个计算可以通过对矩阵 $\begin{bmatrix} b_1 & b_2 & b_3 & p_1 & p_2 \end{bmatrix}$ 做行初等变换化为简化阶梯形而得：

$$\begin{bmatrix} 4 & 0 & 5 & 2 & 1 \\ 0 & 4 & 2 & 0 & 2 \\ 3 & 2 & 4 & 0 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & -2 & \frac{2}{3} \\ 0 & 1 & 0 & -1 & \frac{2}{3} \\ 0 & 0 & 1 & 2 & -\frac{1}{3} \end{bmatrix}$$

取第 4 列得到 p_1，取第 5 列得到 p_2：

$$p_1 = -2b_1 - b_2 + 2b_3 \text{ 和 } p_2 = \frac{2}{3}b_1 + \frac{2}{3}b_2 - \frac{1}{3}b_3$$

p_1 中线性组合的权重和是-1，不是 1. 因此，p_1 不是各 b_i 的一个仿射组合. 然而，由于 p_2 中权重和是 1，所以 p_2 是各 b_i 的一个仿射组合. ■

定义 如果对任意实数 t，由 $p, q \in S$ 得出 $(1-t)p + tq \in S$，则集合 S 是**仿射**的.

几何学上，如果两个点在集合中，过这些点的直线在集合中，则集合是仿射的.（如果 S 只包含一个点 p，则通过 p 的直线和 p 只是一个点，一条“退化”的直线.）代数学上，若一个集合 S 是仿射的，则需要 S 中两点的每个仿射组合都属于 S. 显然，这个定义和 S 包含 S 中任意数量的点的仿射组合是等价的.

定理 2 当且仅当 S 中点的每一个仿射组合都属于 S，集合 S 是仿射的，即当且仅当 $S = \text{aff } S$，S 是仿射的.

证明 假定 S 是仿射的，并对 S 中仿射组合点的数目 m 做推导. 当 m 是 1 或 2 时，由仿射集合的定义，S 中 m 点的仿射组合属于 S. 现在，假设 S 中 k 点或少于 k 点的每一个仿射组合属于 S，考虑 $k+1$ 点的组合. 对 $i = 1, \cdots, k+1$，取 S 中 v_i 为例，并令 $y = c_1v_1 + \cdots + c_kv_k + c_{k+1}v_{k+1}$，这

里 $c_1+\cdots+c_{k+1}=1$. 由于所有 c_i 的和是 1，它们中至少有一个不等于 1. 如果需要，对 $\boldsymbol{v}_i$ 和 c_i 重新标记，我们可以认为 $c_{k+1}\neq 1$，令 $t=c_1+\cdots+c_k$，则 $t=1-c_{k+1}\neq 0$，并且

$$\boldsymbol{y}=(1-c_{k+1})(\frac{c_1}{t}\boldsymbol{v}_1+\cdots+\frac{c_k}{t}\boldsymbol{v}_k)+c_{k+1}\boldsymbol{v}_{k+1} \tag{6}$$

由归纳假设，由于系数和是 1，点 $\boldsymbol{z}=(c_1/t)\boldsymbol{v}_1+\cdots+(c_k/t)\boldsymbol{v}_k$ 属于 S. 从而（6）显示了 $\boldsymbol{y}$ 作为 S 中两点的一个仿射组合是属于 S 的. 由数学归纳法原则，这些点的每一个仿射组合属于 S，即 $\text{aff } S\subset S$. 但反过来，$S\subset \text{aff } S$ 总是成立的. 因此，当 S 为仿射时，$S=\text{aff } S$. 相反，如果 $S=\text{aff } S$，则 S 中两点（或更多点）的仿射组合属于 S，从而 S 是仿射的. ■

下面的定义提供了仿射集合的术语，它暗示了其与 $\mathbb{R}^n$ 子空间的密切关系.

定义 $\mathbb{R}^n$ 中一个集合 S 被向量 $\boldsymbol{p}$ 平移后就是集合 $S+\boldsymbol{p}=\{\boldsymbol{s}+\boldsymbol{p}:\boldsymbol{s}\in S\}$ [㊀]. $\mathbb{R}^n$ 中一个**平面**就是 $\mathbb{R}^n$ 子空间的一个平移. 如果一个平面是另一个平面的平移，则两个平面是**平行**的. 一个**平面的维数**是对应平行子空间的维数. 一个**集合的维数**记为 $\dim S$，是包含 S 的最小平面的维数. $\mathbb{R}^n$ 中一条**直线**是维数为 1 的平面，$\mathbb{R}^n$ 中一个**超平面**是维数为 n-1 的平面.

在 $\mathbb{R}^3$ 中，真子空间[㊁] 包含了原点 $\mathbf{0}$、穿过原点 $\mathbf{0}$ 的所有直线和穿过原点 $\mathbf{0}$ 的所有平面. 因此，$\mathbb{R}^3$ 中真子平面就是点（0 维）、直线（1 维）和平面（2 维），它们可能穿过原点，也可能不穿过原点.

下一个定理表明了 $\mathbb{R}^3$ 中这些直线和平面的集合描述，正好与前面代数描述中两个或三个点的仿射组合是一致的.

定理 3 当且仅当 S 是一个平面时，一个非空集合 S 是仿射的.

证明 假定 S 是仿射的，令 $\boldsymbol{p}$ 是 S 中任意固定点，并且 $W=S+(-\boldsymbol{p})$，从而 $S=W+\boldsymbol{p}$. 为了证明 S 是一个平面，就要证明 W 是 $\mathbb{R}^n$ 的一个子空间. 由于 $\boldsymbol{p}$ 在 S 中，零向量在 W 中. 为了证明 W 对向量的加法和标量乘法运算封闭，就要证明，如果 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 是 W 中的元素，那么对每一个实数 t，$\boldsymbol{u}_1+t\boldsymbol{u}_2$ 属于 W. 由于 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 位于 W 中，存在 S 中 $\boldsymbol{s}_1$ 和 $\boldsymbol{s}_2$，使得 $\boldsymbol{u}_1=\boldsymbol{s}_1-\boldsymbol{p}$ 和 $\boldsymbol{u}_2=\boldsymbol{s}_2-\boldsymbol{p}$. 因此，对每一个实数 t，

$$\begin{aligned}\boldsymbol{u}_1+t\boldsymbol{u}_2&=(\boldsymbol{s}_1-\boldsymbol{p})+t(\boldsymbol{s}_2-\boldsymbol{p})\\&=(1-t)\boldsymbol{s}_1+t(\boldsymbol{s}_1+\boldsymbol{s}_2-\boldsymbol{p})-\boldsymbol{p}\end{aligned}$$

令 $\boldsymbol{y}=\boldsymbol{s}_1+\boldsymbol{s}_2-\boldsymbol{p}$，则 $\boldsymbol{y}$ 是 S 中点的一个仿射组合. 由于 S 是仿射的，$\boldsymbol{y}$ 在 S 中（定理 2）. 但 $(1-t)\boldsymbol{s}_1+t\boldsymbol{y}$ 也在 S 中，所以 $\boldsymbol{u}_1+t\boldsymbol{u}_2$ 在 $-\boldsymbol{p}+S=W$ 中. 这就表明 W 是 $\mathbb{R}^n$ 的一个子空间. 从而由 $S=W+\boldsymbol{p}$，S 是一个平面.

相反，假定 S 是一个平面，即对某些 $\boldsymbol{p}\in\mathbb{R}^n$ 和子空间 W，有 $S=W+\boldsymbol{p}$. 为了表明 S 是仿射的，就要证明对 S 中每一对 $\boldsymbol{s}_1$ 和 $\boldsymbol{s}_2$，穿过 $\boldsymbol{s}_1$ 和 $\boldsymbol{s}_2$ 的直线属于 S. 由 W 的定义，存在 W 中 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 使得 $\boldsymbol{s}_1=\boldsymbol{u}_1+\boldsymbol{p}$ 和 $\boldsymbol{s}_2=\boldsymbol{u}_2+\boldsymbol{p}$. 所以，对每一个实数 t，

$$\begin{aligned}(1-t)\boldsymbol{s}_1+t\boldsymbol{s}_2&=(1-t)(\boldsymbol{u}_1+\boldsymbol{p})+t(\boldsymbol{u}_2+\boldsymbol{p})\\&=(1-t)\boldsymbol{u}_1+t\boldsymbol{u}_2+\boldsymbol{p}\end{aligned}$$

由于 W 是一个子空间，$(1-t)\boldsymbol{u}_1+t\boldsymbol{u}_2\in W$，从而 $(1-t)\boldsymbol{s}_1+t\boldsymbol{s}_2\in W+\boldsymbol{p}=S$. 因此，$S$ 是仿射. ■

㊀ 若 p=0，则 S 的平移是 S 本身. 参见 1.5 节图 1-24.

㊁ A 是 B 的子集，若 $A\neq B$，则 A 称为 B 的真子集. 同样的条件应用于 $\mathbb{R}^n$ 的真子空间和真子平面：它们都不等于 $\mathbb{R}^n$.

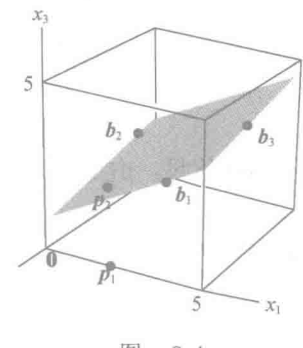


图　8-4

定理 3 给出了一个集合的仿射包的几何解释：由一个集合中的点的所有仿射组合构成的点集是平面. 比如，图 8-4 显示了例 2 中研究的点. 虽然所有 $\boldsymbol{b}_1,\boldsymbol{b}_2$ 和 $\boldsymbol{b}_3$ 的线性组合的集合是 $\mathbb{R}^3$，但所有仿射组合的集合只是穿过 $\boldsymbol{b}_1,\boldsymbol{b}_2$ 和 $\boldsymbol{b}_3$ 的平面. 注意 $\boldsymbol{p}_2$（例 2）在穿过 $\boldsymbol{b}_1,\boldsymbol{b}_2$ 和 $\boldsymbol{b}_3$ 的平面中，而 $\boldsymbol{p}_1$ 不在这个平面中. 还可参见习题 14.

下一个例题从一个新颖的角度来观察一个熟悉的集合——方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 的所有解的集合.

例 3　假定一个方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的解是所有 $\boldsymbol{x}=x_3\boldsymbol{u}+\boldsymbol{p}$ 的形式，这里 $\boldsymbol{u}=\begin{bmatrix}2\\-3\\1\end{bmatrix}$ 和 $\boldsymbol{p}=\begin{bmatrix}4\\0\\-3\end{bmatrix}$. 回忆 1.5 节，这个集合平行于 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集，该解集由形如 $x_3\boldsymbol{u}$ 的点组成. 求点 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 使得 $A\boldsymbol{x}=\boldsymbol{b}$ 的解集是 $\text{aff}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$.

解　解集就是沿着 $\boldsymbol{u}$ 的方向穿过 $\boldsymbol{p}$ 的直线，如图 8-2 所示. 由于 $\text{aff}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是穿过 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的直线，确定直线 $\boldsymbol{x}=x_3\boldsymbol{u}+\boldsymbol{p}$ 上的两个点. 当 $x_3=0$ 和 $x_3=1$ 时，可得两种简单的情况. 即选择 $\boldsymbol{v}_1=\boldsymbol{p}$ 和 $\boldsymbol{v}_2=\boldsymbol{u}+\boldsymbol{p}$，所以

$$\boldsymbol{v}_2=\boldsymbol{u}+\boldsymbol{p}=\begin{bmatrix}2\\-3\\1\end{bmatrix}+\begin{bmatrix}4\\0\\-3\end{bmatrix}=\begin{bmatrix}6\\-3\\-2\end{bmatrix}$$

在这种情况下，解集被描述为所有形如 $\boldsymbol{x}=(1-x_3)\begin{bmatrix}4\\0\\-3\end{bmatrix}+x_3\begin{bmatrix}6\\-3\\-2\end{bmatrix}$ 的仿射组合的集合.

在前面，定理 1 描述了仿射组合和线性组合之间的重要关系. 下面的定理将提供另一个关于仿射组合的研究视角，其中 $\mathbb{R}^2$ 和 $\mathbb{R}^3$ 的情况与计算机图形学方面的应用密切相关，这将在下节中讨论（还有 2.7 节）.

定义　对 $\mathbb{R}^n$ 中的 $\boldsymbol{v}$，$\boldsymbol{v}$ 的**标准齐次形式**是 $\mathbb{R}^{n+1}$ 中的点 $\tilde{\boldsymbol{v}}=\begin{bmatrix}\boldsymbol{v}\\1\end{bmatrix}$.

定理 4　$\mathbb{R}^n$ 中的一个点 $\boldsymbol{y}$ 是 $\mathbb{R}^n$ 中 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 的一个仿射组合当且仅当 $\boldsymbol{y}$ 的齐次形式在 $\text{Span}\{\tilde{\boldsymbol{v}}_1,\cdots,\tilde{\boldsymbol{v}}_p\}$ 中，即 $\boldsymbol{y}=c_1\boldsymbol{v}_1+\cdots+c_p\boldsymbol{v}_p$ 且 $c_1+\cdots+c_p=1$，当且仅当 $\tilde{\boldsymbol{y}}=c_1\tilde{\boldsymbol{v}}_1+\cdots+c_p\tilde{\boldsymbol{v}}_p$.

证明　显然，点 $\boldsymbol{y}$ 在 $\text{aff}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 中当且仅当存在权重 $c_1,\cdots,c_p$，使得

$$\begin{bmatrix}\boldsymbol{y}\\1\end{bmatrix}=c_1\begin{bmatrix}\boldsymbol{v}_1\\1\end{bmatrix}+c_2\begin{bmatrix}\boldsymbol{v}_2\\1\end{bmatrix}+\cdots+c_p\begin{bmatrix}\boldsymbol{v}_p\\1\end{bmatrix}$$

而这种情况发生当且仅当 $\tilde{\boldsymbol{y}}$ 在 $\text{Span}\{\tilde{\boldsymbol{v}}_1,\tilde{\boldsymbol{v}}_2,\cdots,\tilde{\boldsymbol{v}}_p\}$ 中. 定理得到了证明. ■

例 4　令 $\boldsymbol{v}_1=\begin{bmatrix}3\\1\\1\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}1\\2\\2\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}1\\7\\1\end{bmatrix},\boldsymbol{p}=\begin{bmatrix}4\\3\\0\end{bmatrix}$. 如果可能，用定理 4 把 $\boldsymbol{p}=\begin{bmatrix}4\\3\\0\end{bmatrix}$ 写成 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 的

仿射组合.

解 对下面方程的增广矩阵进行行化简：

$$x_1\tilde{\boldsymbol{v}}_1 + x_2\tilde{\boldsymbol{v}}_2 + x_3\tilde{\boldsymbol{v}}_3 = \tilde{\boldsymbol{p}}$$

为了简化算法，把由 1 组成的第四行移到第一行（等价于做三个行对换），然后把矩阵化为简化阶梯形如下：

$$\begin{bmatrix}\tilde{\boldsymbol{v}}_1 & \tilde{\boldsymbol{v}}_2 & \tilde{\boldsymbol{v}}_3 & \tilde{\boldsymbol{p}}\end{bmatrix} \sim \begin{bmatrix}1&1&1&1\\3&1&1&4\\1&2&7&3\\1&2&1&0\end{bmatrix} \sim \begin{bmatrix}1&1&1&1\\0&-2&-2&1\\0&1&6&2\\0&1&0&-1\end{bmatrix}$$

$$\sim \cdots \sim \begin{bmatrix}1&0&0&1.5\\0&1&0&-1\\0&0&1&0.5\\0&0&0&0\end{bmatrix}$$

由定理 4，得 $1.5\boldsymbol{v}_1 - \boldsymbol{v}_2 + 0.5\boldsymbol{v}_3 = \boldsymbol{p}$，见图 8-5，图中显示了含有 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ 和 $\boldsymbol{p}$（以及坐标轴上的点）的平面.

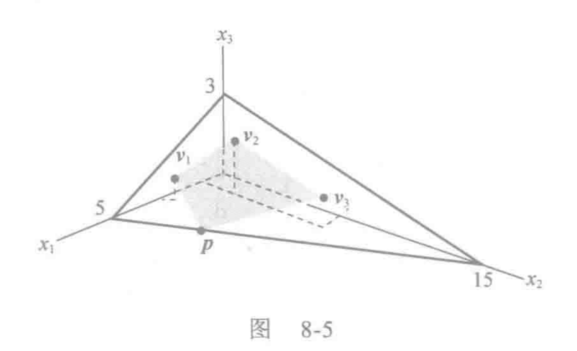


图 8-5

练习题

画出 $\boldsymbol{v}_1 = \begin{bmatrix}1\\0\end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix}-1\\2\end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix}3\\1\end{bmatrix}, \boldsymbol{p} = \begin{bmatrix}4\\3\end{bmatrix}$ 的图形，并解释为什么 $\boldsymbol{p}$ 必须是 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ 的仿射组合. 然后求出 $\boldsymbol{p}$ 的仿射组合.

习题 8.1

在习题 1~4 中，如果可能，把 $\boldsymbol{y}$ 写成其他点的仿射组合.

1. $\boldsymbol{v}_1 = \begin{bmatrix}1\\2\end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix}-2\\2\end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix}0\\4\end{bmatrix}, \boldsymbol{v}_4 = \begin{bmatrix}3\\7\end{bmatrix}, \boldsymbol{y} = \begin{bmatrix}5\\3\end{bmatrix}$

2. $\boldsymbol{v}_1 = \begin{bmatrix}1\\1\end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix}-1\\2\end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix}3\\2\end{bmatrix}, \boldsymbol{y} = \begin{bmatrix}5\\7\end{bmatrix}$

3. $\boldsymbol{v}_1 = \begin{bmatrix}-3\\1\\1\end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix}0\\4\\-2\end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix}4\\-2\\6\end{bmatrix}, \boldsymbol{y} = \begin{bmatrix}17\\1\\5\end{bmatrix}$

4. $\boldsymbol{v}_1 = \begin{bmatrix}1\\2\\0\end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix}2\\-6\\7\end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix}4\\3\\1\end{bmatrix}, \boldsymbol{y} = \begin{bmatrix}-3\\4\\-4\end{bmatrix}$

在习题 5 和 6 中，令 $b_1=\begin{bmatrix}2\\1\\1\end{bmatrix},b_2=\begin{bmatrix}1\\0\\-2\end{bmatrix}$，$b_3=\begin{bmatrix}2\\-5\\1\end{bmatrix}$，$S=\{b_1,b_2,b_3\}$，注意 S 是 $\mathbb{R}^3$ 中的一个正交基. 如果可能，把给定点写成 S 中点的仿射组合.（提示：使用 6.2 节的定理 5 而不是行化简来求权重.）.

5. a. $p_1=\begin{bmatrix}3\\8\\4\end{bmatrix}$　b. $p_2=\begin{bmatrix}6\\-3\\3\end{bmatrix}$　c. $p_3=\begin{bmatrix}0\\-1\\-5\end{bmatrix}$

6. a. $p_1=\begin{bmatrix}0\\-19\\-5\end{bmatrix}$　b. $p_2=\begin{bmatrix}1.5\\-1.3\\-0.5\end{bmatrix}$　c. $p_3=\begin{bmatrix}5\\-4\\0\end{bmatrix}$

7. 令 $v_1=\begin{bmatrix}1\\0\\3\\0\end{bmatrix},v_2=\begin{bmatrix}2\\-1\\0\\4\end{bmatrix},v_3=\begin{bmatrix}-1\\2\\1\\1\end{bmatrix}$，$p_1=\begin{bmatrix}5\\-3\\5\\3\end{bmatrix},p_2=\begin{bmatrix}-9\\10\\9\\-13\end{bmatrix}$，$p_3=\begin{bmatrix}4\\2\\8\\5\end{bmatrix}$，并且 $S=\{v_1,v_2,v_3\}$. 可以证明 S 是线性无关的.
   a. p_1 属于 Span S 吗？ p_1 属于 aff S 吗？
   b. p_2 属于 Span S 吗？ p_2 属于 aff S 吗？
   c. p_3 属于 Span S 吗？ p_3 属于 aff S 吗？

8. 根据下列已给向量仿照习题 7 重做.

$$v_1=\begin{bmatrix}1\\0\\3\\-2\end{bmatrix},v_2=\begin{bmatrix}2\\1\\6\\-5\end{bmatrix},v_3=\begin{bmatrix}3\\0\\12\\-6\end{bmatrix},\quad p_1=\begin{bmatrix}4\\-1\\15\\-7\end{bmatrix},p_2=\begin{bmatrix}-5\\3\\-8\\6\end{bmatrix},$$

$$p_3=\begin{bmatrix}1\\6\\-6\\-8\end{bmatrix}$$

9. 假定一个方程 $Ax=b$ 的解是所有 $x=x_3u+p$ 的形式，这里 $u=\begin{bmatrix}4\\-2\end{bmatrix}$ 和 $p=\begin{bmatrix}-3\\0\end{bmatrix}$. 求点 v_1 和 v_2，使得 $Ax=b$ 的解集是 aff$\{v_1,v_2\}$.

10. 假定一个方程 $Ax=b$ 的解是所有 $x=x_3u+p$ 的形式，这里

$$u=\begin{bmatrix}5\\1\\-2\end{bmatrix}\text{ 和 }p=\begin{bmatrix}1\\-3\\4\end{bmatrix}$$

求出点 v_1 和 v_2，使得 $Ax=b$ 的解集是 aff$\{v_1,v_2\}$.

在习题 11 和 12 中，判断每个叙述的正误，并说明理由.

11. a. 集合 S 中点的所有仿射组合称为 S 的仿射包.
   b. 如果 $\{b_1,\cdots,b_k\}$ 是 $\mathbb{R}^n$ 的线性独立子集，并且如果 p 是 $b_1,\cdots,b_k$ 的线性组合，那么 p 是 $b_1,\cdots,b_k$ 的一个仿射组合.
   c. 两个不同点的仿射包称为一条直线.
   d. 一个平面是一个子空间.
   e. $\mathbb{R}^3$ 中一个平面是一个超平面.

12. a. 如果 $S=\{x\}$，那么 aff S 是一个空集.
   b. 当且仅当一个集合包含它的仿射包时，它是一个仿射集.
   c. 一个一维的平面称为一条直线.
   d. 一个二维的平面称为一个超平面.
   e. 穿过原点的平面是子空间.

13. 假定 $\{v_1,v_2,v_3\}$ 是 $\mathbb{R}^3$ 中的一个基. 证明：$\{v_2-v_1,v_3-v_1\}$ 是 $\mathbb{R}^3$ 中一个平面.（提示：当 Span$\{u,v\}$是一个平面时，关于 u 和 v，你能得出什么结论？）

14. 证明：如果 $\{v_1,v_2,v_3\}$ 是 $\mathbb{R}^3$ 的一个基，则 aff$\{v_1,v_2,v_3\}$ 是穿过 v_1,v_2 和 v_3 的平面.

15. 令 A 是一个 $m\times n$ 矩阵，给定 $\mathbb{R}^m$ 中 b，证明：$Ax=b$ 的所有解的集合 S 是 $\mathbb{R}^n$ 的一个仿射子集.

16. 令 $v\in\mathbb{R}^n$ 和 $k\in\mathbb{R}$. 证明：$S=\left\{x\in\mathbb{R}^n:x\cdot v=k\right\}$ 是 $\mathbb{R}^n$ 的一个仿射子集.

17. 选择三个点的集合 S，使得 aff S 是 $\mathbb{R}^3$ 中的平面，它的方程是 $x_3=5$，并证明.

18. 选择 $\mathbb{R}^3$ 中 4 个不同点的集合，使得 aff S 是平

面 $2x_1+x_2-3x_3=12$，并证明.

19. 令 S 是 $\mathbb{R}^n$ 的一个仿射子集，假定 $f:\mathbb{R}^n\to\mathbb{R}^m$ 是一个线性变换，并令 $f(S)$ 表示图形 $\{f(x):x\in S\}$ 的集合，证明：$f(S)$ 是 $\mathbb{R}^m$ 的一个仿射子集.
20. 令 $f:\mathbb{R}^n\to\mathbb{R}^m$ 是一个线性变换，T 是 $\mathbb{R}^m$ 的仿射子集，$S=\{\boldsymbol{x}\in\mathbb{R}^n:f(\boldsymbol{x})\in T\}$. 证明：$S$ 是 $\mathbb{R}^n$ 的一个仿射子集.

在习题 21~26 中，证明给定的关于 $\mathbb{R}^n$ 中子集 A 和 B 的陈述，或者提供 $\mathbb{R}^2$ 中所要求的例子. 证明可以用前面习题的结论（也可以用本书中已学过的定理）.

21. 如果 $A\subset B$ 且 B 是仿射的，那么 $\text{aff}\ A\subset B$.
22. 如果 $A\subset B$，那么 $\text{aff}\ A\subset\text{aff}\ B$.
23. $[(\text{aff}\ A)\cup(\text{aff}\ B)]\subset\text{aff}(A\cup B)$.（提示：为了证明 $D\cup E\subset F$，证明 $D\subset F$ 且 $E\subset F$.）
24. 在 $\mathbb{R}^2$ 中寻找一个例子来证明习题 23 中的等式不成立.（提示：考虑集合 A 和 B，其中每个集合仅包含一个或两个点.）
25. $\text{aff}(A\cap B)\subset(\text{aff}\ A\cap\text{aff}\ B)$.
26. 在 $\mathbb{R}^2$ 中寻找一个例子来证明习题 25 中的等式不成立.

练习题答案

由于点 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 和 $\boldsymbol{v}_3$ 不在同一条直线上，$\text{aff}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 不是一维的. 因此，$\text{aff}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 必等于 $\mathbb{R}^2$. 为了求出把 $\boldsymbol{p}$ 表示成 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 和 $\boldsymbol{v}_3$ 的一个仿射组合的权重，首先计算

$$\boldsymbol{v}_2-\boldsymbol{v}_1=\begin{bmatrix}-2\\2\end{bmatrix},\quad \boldsymbol{v}_3-\boldsymbol{v}_1=\begin{bmatrix}2\\1\end{bmatrix},\quad \boldsymbol{p}-\boldsymbol{v}_1=\begin{bmatrix}3\\3\end{bmatrix}$$

为把 $\boldsymbol{p}-\boldsymbol{v}_1$ 写成 $\boldsymbol{v}_2-\boldsymbol{v}_1$ 和 $\boldsymbol{v}_3-\boldsymbol{v}_1$ 的线性组合，对矩阵做行简化运算得到：

$$\begin{bmatrix}-2&2&3\\2&1&3\end{bmatrix}\sim\begin{bmatrix}1&0&\frac{1}{2}\\0&1&2\end{bmatrix}$$

因此，$\boldsymbol{p}-\boldsymbol{v}_1=\frac{1}{2}(\boldsymbol{v}_2-\boldsymbol{v}_1)+2(\boldsymbol{v}_3-\boldsymbol{v}_1)$，这表明

$$\boldsymbol{p}=(1-\frac{1}{2}-2)\boldsymbol{v}_1+\frac{1}{2}\boldsymbol{v}_2+2\boldsymbol{v}_3=-\frac{3}{2}\boldsymbol{v}_1+\frac{1}{2}\boldsymbol{v}_2+2\boldsymbol{v}_3$$

由于系数和为 1，故 $\boldsymbol{p}$ 可以表示成 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 的仿射组合.

另外一种方法，运用例 3 中的方法和行简化运算：

$$\begin{bmatrix}\boldsymbol{v}_1&\boldsymbol{v}_2&\boldsymbol{v}_3&\boldsymbol{p}\\1&1&1&1\end{bmatrix}\sim\begin{bmatrix}1&1&1&1\\1&-1&3&4\\0&2&1&3\end{bmatrix}\sim\begin{bmatrix}1&0&0&-\frac{3}{2}\\0&1&0&\frac{1}{2}\\0&0&1&2\end{bmatrix}$$

这就表明了 $\boldsymbol{p}=-\frac{3}{2}\boldsymbol{v}_1+\frac{1}{2}\boldsymbol{v}_2+2\boldsymbol{v}_3$.

8.2 仿射无关性

本节继续探讨线性概念和仿射概念之间的关系. 考虑 $\mathbb{R}^3$ 中三个向量的集合，记为 $S=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$. 如果 S 是线性相关的，那么其中一个向量是其他两个向量的线性组合. 当一个向量是其他两个向量的仿射组合时，又会怎么样呢？比如，假定对某些 $\mathbb{R}$ 中的 t，有

$$\boldsymbol{v}_3=(1-t)\boldsymbol{v}_1+t\boldsymbol{v}_2$$

则

$$(1-t)\boldsymbol{v}_1+t\boldsymbol{v}_2-\boldsymbol{v}_3=\boldsymbol{0}$$

由于不是所有的权重都等于 0，因此这是一个线性相关的关系. 而且，线性相关关系中的权重和为 0：

$$(1-t)+t+(-1)=0$$

这也是定义仿射相关的附加特性.

仿射无关性概念

定义 设 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 是 $\mathbb{R}^n$ 中的一个指标点集（译者注：指标点集指的是集合中指标不同的元素可以表示同一点），如果存在不全为零的实数 $c_1,\cdots,c_p$，使得

$$c_1+\cdots+c_p=0\text{，及 } c_1\boldsymbol{v}_1+\cdots+c_p\boldsymbol{v}_p=\boldsymbol{0} \tag{1}$$

则称指标点集 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 是**仿射相关**的. 否则，称该集合是**仿射无关**的.

仿射组合是线性组合的一种特殊情形，并且仿射相关是附带有一个限制条件的线性相关. 因此，每一个仿射相关集都是自动线性相关的.

只有一个点的集合 $\{\boldsymbol{v}_1\}$ 必是仿射无关的，因为当只有一个系数时，所需系数 c_i 的特性不能得到满足. 事实上，对 $\{\boldsymbol{v}_1\}$ 来说，（1）中的第一个方程正是 $c_1=0$，另一方面又要求 c_1 不为零.

习题 13 要求证明当且仅当 $\boldsymbol{v}_1=\boldsymbol{v}_2$ 时，指标集合 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是仿射相关的. 下面的定理解决了一般的情形，并证明仿射相关的概念与线性相关概念的相似性.（c）和（d）给出了确定一个集合是否是仿射相关的有用方法. 回忆 8.1 节，如果 $\boldsymbol{v}$ 在 $\mathbb{R}^n$ 中，那么 $\mathbb{R}^{n+1}$ 中向量 $\tilde{\boldsymbol{v}}$ 表示 $\boldsymbol{v}$ 的齐次形式.

定理 5 给定 $\mathbb{R}^n$ 中的一个指标集合 $S=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$，$p\geqslant 2$，下面的叙述是逻辑等价的. 也就是说，它们同真，或者同假.

a. S 是仿射相关的.

b. S 中有一个点是 S 中其他点的一个仿射组合.

c. $\mathbb{R}^n$ 中集合 $\{\boldsymbol{v}_2-\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p-\boldsymbol{v}_1\}$ 是线性相关的.

d. $\mathbb{R}^{n+1}$ 中集合 $\{\tilde{\boldsymbol{v}}_1,\cdots,\tilde{\boldsymbol{v}}_p\}$（齐次形式）是线性相关的.

证明 假定（a）为真，令 $c_1,\cdots,c_p$ 满足（1）. 如果需要，对这些点重新命名，我们可以认为 $c_1\neq 0$，并将（1）中的方程都除以 c_1，从而，$1+(c_2/c_1)+\cdots+(c_p/c_1)=0$，并且

$$\boldsymbol{v}_1=(-c_2/c_1)\boldsymbol{v}_2+\cdots+(-c_p/c_1)\boldsymbol{v}_p \tag{2}$$

注意，（2）中右边的系数和为 1. 因此由（a）得出（b）. 现在，假定（b）为真，如果需要，对点重新命名，可以认为 $\boldsymbol{v}_1=c_2\boldsymbol{v}_2+\cdots+c_p\boldsymbol{v}_p$，这里 $c_2+\cdots+c_p=1$. 则

$$(c_2+\cdots+c_p)\boldsymbol{v}_1=c_2\boldsymbol{v}_2+\cdots+c_p\boldsymbol{v}_p \tag{3}$$

从而

$$c_2(\boldsymbol{v}_2-\boldsymbol{v}_1)+\cdots+c_p(\boldsymbol{v}_p-\boldsymbol{v}_1)=\boldsymbol{0} \tag{4}$$

由于它们的和为 1，所以并不是所有的 $c_2,\cdots,c_p$ 都是 0. 因此，由（b）可得出（c）.

接下来，如果（c）为真，那么存在不全为零的权重 $c_2,\cdots,c_p$，使得（4）成立. 把（4）重写为（3）的形式，并且 $c_1=-(c_2+\cdots+c_p)$. 则 $c_1+\cdots+c_p=0$. 因此，由（3）可证明（1）成立. 所

以由（c）得出（a），这就证明了（a）、（b）和（c）是逻辑等价的. 最后，由于（1）中的两个方程等价于下面的包含 S 中点的齐次形式的方程，所以（d）等价于（a）：

$$c_1\begin{bmatrix}\boldsymbol{v}_1\\1\end{bmatrix}+\cdots+c_p\begin{bmatrix}\boldsymbol{v}_p\\1\end{bmatrix}=\begin{bmatrix}\boldsymbol{0}\\0\end{bmatrix}$$ ■

在定理 5（c）中，$\boldsymbol{v}_1$ 可以被 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 中任何其他的点所取代. 只需要改变证明中的标记就行了. 因此，为了检测一个集合是否是仿射相关的，可以从集合中的其他点中减去一个点，并检验转化后的 $p-1$ 点集是否是线性相关的.

例 1 两个不同点 $\boldsymbol{p}$ 和 $\boldsymbol{q}$ 的仿射包是一条直线. 如果第三个点 $\boldsymbol{r}$ 在直线上，那么 $\{\boldsymbol{p},\boldsymbol{q},\boldsymbol{r}\}$ 是一个仿射相关集. 如果一个点 $\boldsymbol{s}$ 不在穿过 $\boldsymbol{p}$ 和 $\boldsymbol{q}$ 的直线上，那么这三点不在同一条直线上，从而 $\{\boldsymbol{p},\boldsymbol{q},\boldsymbol{s}\}$ 是仿射无关集，见图 8-6.

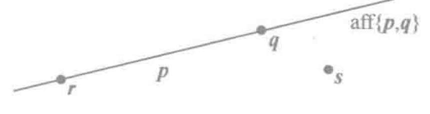


图 8-6 $\{\boldsymbol{p},\boldsymbol{q},\boldsymbol{r}\}$ 是仿射相关的

例 2 令 $\boldsymbol{v}_1=\begin{bmatrix}1\\3\\7\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}2\\7\\6.5\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}0\\4\\7\end{bmatrix}$，$S=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$，确定 S 是否是仿射无关的.

解 计算 $\boldsymbol{v}_2-\boldsymbol{v}_1=\begin{bmatrix}1\\4\\-0.5\end{bmatrix}$ 和 $\boldsymbol{v}_3-\boldsymbol{v}_1=\begin{bmatrix}-1\\1\\0\end{bmatrix}$. 这两个点不成倍数，因此，它们构成了一个线性无关集 S'. 从而定理 5 中的所有陈述都是错误的，并且 S 是仿射无关的. 图 8-7 表明了 S 和平移集 S'. 注意 Span S' 是穿过原点的一个平面，aff S 是与之平行并通过 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 的平面.（当然，这里仅显示了每个平面的一部分.）

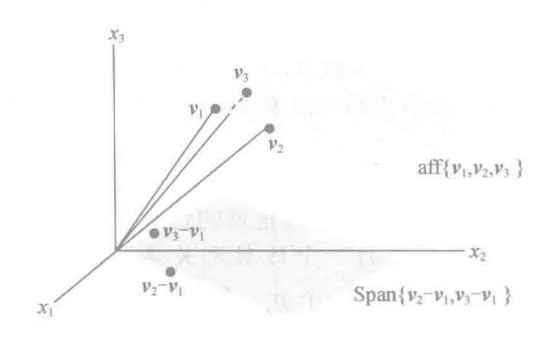


图 8-7 一个仿射无关集 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$

例 3 令 $\boldsymbol{v}_1=\begin{bmatrix}1\\3\\7\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}2\\7\\6.5\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}0\\0.4\\7\end{bmatrix}$，$\boldsymbol{v}_4=\begin{bmatrix}0\\14\\6\end{bmatrix}$，并令 $S=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_4\}$. S 是仿射相关的吗？

解 计算 $\boldsymbol{v}_2-\boldsymbol{v}_1=\begin{bmatrix}1\\4\\-0.5\end{bmatrix},\boldsymbol{v}_3-\boldsymbol{v}_1=\begin{bmatrix}-1\\1\\0\end{bmatrix}$，$\boldsymbol{v}_4-\boldsymbol{v}_1=\begin{bmatrix}-1\\11\\-1\end{bmatrix}$，并对矩阵进行行化简运算：

$$\begin{bmatrix} 1 & -1 & -1 \\ 4 & 1 & 11 \\ -0.5 & 0 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & -1 \\ 0 & 5 & 15 \\ 0 & -0.5 & -1.5 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & -1 \\ 0 & 5 & 15 \\ 0 & 0 & 0 \end{bmatrix}$$

回忆 4.6 节（或 2.8 节），由于并不是每一列都是一个主元列，因此这些列是线性相关的. 所以，$\boldsymbol{v}_2-\boldsymbol{v}_1, \boldsymbol{v}_3-\boldsymbol{v}_1$ 和 $\boldsymbol{v}_4-\boldsymbol{v}_1$ 是线性相关的. 由定理 5 中的（c），$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4\}$ 是仿射相关的. 这个相关性也可以用定理 5（d）取代（c）而得到. ■

例 3 的计算表明 $\boldsymbol{v}_4-\boldsymbol{v}_1$ 是 $\boldsymbol{v}_2-\boldsymbol{v}_1$ 和 $\boldsymbol{v}_3-\boldsymbol{v}_1$ 的线性组合，这说明 $\boldsymbol{v}_4-\boldsymbol{v}_1$ 在 $\text{Span}\{\boldsymbol{v}_2-\boldsymbol{v}_1, \boldsymbol{v}_3-\boldsymbol{v}_1\}$ 中. 由 8.1 节定理 1，$\boldsymbol{v}_4$ 在 $\text{aff}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 中. 实际上，例 3 中矩阵进一步的行化简运算将表明

$$\boldsymbol{v}_4-\boldsymbol{v}_1 = 2(\boldsymbol{v}_2-\boldsymbol{v}_1)+3(\boldsymbol{v}_3-\boldsymbol{v}_1) \tag{5}$$

$$\boldsymbol{v}_4 = -4\boldsymbol{v}_1+2\boldsymbol{v}_2+3\boldsymbol{v}_3 \tag{6}$$

见图 8-8.

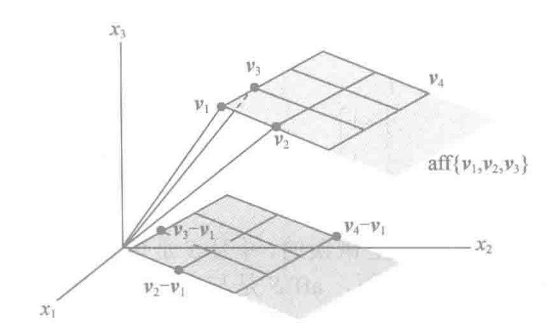


图 8-8 $\boldsymbol{v}_4$ 在平面 $\text{aff}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 中

图 8-8 显示了 $\text{Span}(\boldsymbol{v}_2-\boldsymbol{v}_1, \boldsymbol{v}_3-\boldsymbol{v}_1)$ 和 $\text{aff}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 的网格. $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 上的网格是基于（5）的. 另一个坐标系统是基于（6）的，其中系数−4,2 和 3 称为 $\boldsymbol{v}_4$ 的**仿射**或**重心坐标**.

重心坐标

重心坐标的定义是基于 4.4 节中表达唯一性定理的仿射形式，见本节习题 17 中的证明.

定理 6 令 $S=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k\}$ 是 $\mathbb{R}^n$ 中的一个仿射无关集. 则 aff S 中每一个 $\boldsymbol{p}$ 都有唯一的 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k$ 的仿射组合表示. 也就是说，对每一个 $\boldsymbol{p}$，存在唯一的标量集 $c_1,\cdots,c_k$，使得

$$\boldsymbol{p}=c_1\boldsymbol{v}_1+\cdots+c_k\boldsymbol{v}_k \text{ 且 } c_1+\cdots+c_k=1 \tag{7}$$

定义 令 $S=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k\}$ 是一个仿射无关集. 则对 aff S 中每一个点 $\boldsymbol{p}$，（7）中 $\boldsymbol{p}$ 的唯一表达式中的系数 $c_1,\cdots,c_p$ 称为 $\boldsymbol{p}$ 的**重心坐标**(或者称为**仿射坐标**).

观察（7）得，它等价于一单个方程

$$\begin{bmatrix} \boldsymbol{p} \\ 1 \end{bmatrix} = c_1\begin{bmatrix} \boldsymbol{v}_1 \\ 1 \end{bmatrix}+\cdots+c_k\begin{bmatrix} \boldsymbol{v}_k \\ 1 \end{bmatrix} \tag{8}$$

包含所有点的齐次形式. 对（8）中的增广矩阵 $[\tilde{\boldsymbol{v}}_1 \quad \cdots \quad \tilde{\boldsymbol{v}}_k \quad \boldsymbol{p}]$ 做行化简运算可得到 $\boldsymbol{p}$ 的重心坐标.

例 4 令 $\boldsymbol{a}=\begin{bmatrix}1\\7\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}3\\0\end{bmatrix},\boldsymbol{c}=\begin{bmatrix}9\\3\end{bmatrix},\ \boldsymbol{p}=\begin{bmatrix}5\\3\end{bmatrix}$. 求由一个仿射集 $\{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}\}$ 决定的 $\boldsymbol{p}$ 的重心坐标.

解 对点的齐次形式下形如（8）的方程的增广矩阵做行简化运算，把最后一行移到第一行后简化运算：

$$[\tilde{\boldsymbol{a}}\quad \tilde{\boldsymbol{b}}\quad \tilde{\boldsymbol{c}}\quad \tilde{\boldsymbol{p}}]=\begin{bmatrix}1&3&9&5\\7&0&3&3\\1&1&1&1\end{bmatrix}\sim\begin{bmatrix}1&1&1&1\\1&3&9&5\\7&0&3&3\end{bmatrix}\sim\begin{bmatrix}1&0&0&\frac{1}{4}\\0&1&0&\frac{1}{3}\\0&0&1&\frac{5}{12}\end{bmatrix}$$

得坐标为 $\frac{1}{4},\frac{1}{3}$ 和 $\frac{5}{12}$，从而 $\boldsymbol{p}=\frac{1}{4}\boldsymbol{a}+\frac{1}{3}\boldsymbol{b}+\frac{5}{12}\boldsymbol{c}$. ■

重心坐标有物理和几何学上的解释. 它最初由 A. F. Moebius 在 1827 年定义，当时这个定义是对三个顶点为 $\boldsymbol{a},\boldsymbol{b}$ 和 $\boldsymbol{c}$ 的三角形中的一个点 $\boldsymbol{p}$ 而言. 他写到，$\boldsymbol{p}$ 的重心坐标就是三个非负的数 m_a,m_b 和 m_c，使得 $\boldsymbol{p}$ 是由三角形（无质量）和三角形顶点处分别放置质量为 m_a,m_b 和 m_c 的质点构成的质量体系的质心. 这些质点的质量可以通过要求它们的和为 1 来唯一确定. 这个观点在物理中至今仍然很有用. ㊀

图 8-9 给出了例 4 中重心坐标的几何解释，显示了三角形 $\triangle\boldsymbol{abc}$ 和它的三个小三角形 $\triangle\boldsymbol{pbc},\triangle\boldsymbol{apc}$ 和 $\triangle\boldsymbol{abp}$. 三个小三角形的面积与 $\boldsymbol{p}$ 的重心坐标成正比，实际上，

$$\begin{aligned}&\text{面积}(\triangle\boldsymbol{pbc})=\frac{1}{4}\cdot\text{面积}(\triangle\boldsymbol{abc})\\&\text{面积}(\triangle\boldsymbol{apc})=\frac{1}{3}\cdot\text{面积}(\triangle\boldsymbol{abc})\\&\text{面积}(\triangle\boldsymbol{abp})=\frac{5}{12}\cdot\text{面积}(\triangle\boldsymbol{abc})\end{aligned}\tag{9}$$

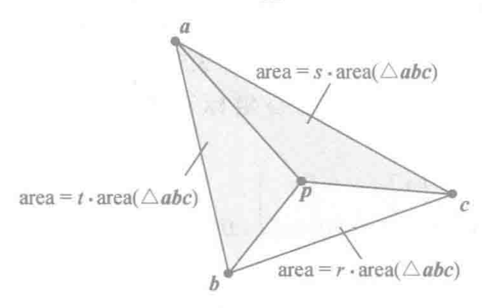


图 8-9 $\boldsymbol{p}=r\boldsymbol{a}+s\boldsymbol{b}+t\boldsymbol{c}$. 其中 $r=\frac{1}{4}$，$s=\frac{1}{3}$，$t=\frac{5}{12}$

图 8-9 中的公式在习题 21~23 会给出验证. 当 $\boldsymbol{p}$ 是 $\mathbb{R}^3$ 中一个四面体内的一点，对应的顶点是 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 和 $\boldsymbol{d}$ 时，类似的性质也成立.

当一个点不在三角形（或四面体）中时，重心的某些或全部坐标值将是负的. 对上述的顶

㊀ 参见 1.3 节习题 29. 然而在天文学中，“重心坐标”通常指普通的 $\mathbb{R}^3$ 点坐标，现在称为国际天球参考系统，外层空间的笛卡尔坐标系统，其原点位于太阳系的质量中心（重心）.

点 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 和坐标值 r,s,t，图 8-10 显示了一个三角形的情形. 比如，穿过 $\boldsymbol{b}$ 和 $\boldsymbol{c}$ 的直线上的点，有 $r=0$，因为它们仅是 $\boldsymbol{b}$ 和 $\boldsymbol{c}$ 的仿射组合. 穿过 $\boldsymbol{a}$ 且平行于上述直线的直线上的点有 $r=1$.

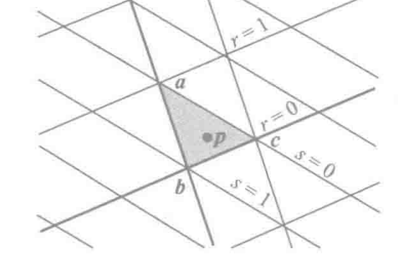


图 8-10 aff{$\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$}中点的重心坐标

计算机图形的重心坐标

当用计算机画图程序来处理几何对象时，设计者会用“线框”逼近对象中的一些关键点来生成和处理对象的图形[㊀]. 比如，如果物体的一部分表面是由一些小三角平面组成的，那么在仅仅知道这些三角面的顶点处的信息的情况下，图形程序就可以很容易地对每一个小三角面进行加色、光及着色. 重心坐标提供了一个由顶点信息平滑地插入三角面内部点信息的工具. 内部点的参数值是顶点处参数值的一个简单的线性组合，组合的权重为该点的重心坐标.

计算机屏幕上的颜色通常可以由 RGB 坐标给出. 一个三参数组 (r,g,b) 表示各颜色（红、绿和蓝）的数量，参数变化范围从 0 到 1. 比如，纯红是 (1,0,0)，白色是 (1,1,1)，黑色是 (0,0,0).

例 5 令 $\boldsymbol{v}_1=\begin{bmatrix}3\\1\\5\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}4\\3\\4\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}1\\5\\1\end{bmatrix}$，$\boldsymbol{p}=\begin{bmatrix}3\\3\\3.5\end{bmatrix}$. 在一个三角形的顶点 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 和 $\boldsymbol{v}_3$ 处的颜色分别是品红 (1,0,1)、浅品红 (1,0.4,1) 和紫色 (0.6,0,1). 求出 $\boldsymbol{p}$ 点的插入颜色，见图 8-11.

图 8-11 插入颜色

解 首先，求出 $\boldsymbol{p}$ 点的重心坐标，这里运用点的齐次形式，第一步先把第 4 行移到第 1 行：

$$\begin{bmatrix}\tilde{\boldsymbol{v}}_1 & \tilde{\boldsymbol{v}}_2 & \tilde{\boldsymbol{v}}_3 & \tilde{\boldsymbol{p}}\end{bmatrix}\sim\begin{bmatrix}1&1&1&1\\3&4&1&3\\1&3&5&3\\5&4&1&3.5\end{bmatrix}\sim\begin{bmatrix}1&0&0&0.25\\0&1&0&0.50\\0&0&1&0.25\\0&0&0&0\end{bmatrix}$$

从而 $\boldsymbol{p}=0.25\boldsymbol{v}_1+0.5\boldsymbol{v}_2+0.25\boldsymbol{v}_3$. 运用 $\boldsymbol{p}$ 的重心坐标对颜色数据做一个线性组合. $\boldsymbol{p}$ 的 RGB 值是：

$$0.25\begin{bmatrix}1\\0\\1\end{bmatrix}+0.50\begin{bmatrix}1\\0.4\\1\end{bmatrix}+0.25\begin{bmatrix}0.6\\0\\1\end{bmatrix}=\begin{bmatrix}0.9\\0.2\\1\end{bmatrix}\begin{matrix}\text{红}\\\text{绿}\\\text{蓝}\end{matrix}$$

■

把图形显示在计算机屏幕上的最后一步是消除“隐藏面”，这些隐藏面都不会在屏幕上看到. 设想屏幕是由 100 万像素的点组成，并考虑一条射线或“光线”是从一个观察者的眼睛，通过一个像素点到达所要 3D 显示的物体上. 屏幕上该像素的颜色和其他信息可以从与射线第一次相交的物体上得到，见图 8-12. 当图像中的物体由线框和三角面逼近时，隐藏面问题可运用重心坐标来解决.

㊀ 第 2 章的介绍性实例展示了波音 777 飞机的线形轮廓模型，用于显现飞机表面的气流.

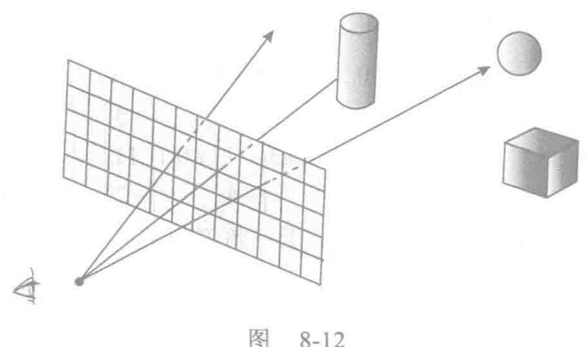


图 8-12

射线-三角面相交的数学方法也可以很好地应用于对物体进行逼真的着色. 目前，这个光线跟踪着色法对实时渲染太慢，但现在硬件的迅速发展将改善这一状况. ㊀

例 6 令 $\boldsymbol{v}_1=\begin{bmatrix}1\\1\\-6\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}8\\1\\-4\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}5\\11\\-2\end{bmatrix},\boldsymbol{a}=\begin{bmatrix}0\\0\\10\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}0.7\\0.4\\-3\end{bmatrix}$，并且对 $t\geqslant 0$ 有 $\boldsymbol{x}(t)=\boldsymbol{a}+t\boldsymbol{b}$，求射线 $\boldsymbol{x}(t)$ 和由三角形三个顶点 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 和 $\boldsymbol{v}_3$ 构成的平面相交的点. 这个点在三角形中吗？

解 平面是 $\mathrm{aff}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$. 这个平面中的一个通用点可写为 $(1-c_2-c_3)\boldsymbol{v}_1+c_2\boldsymbol{v}_2+c_3\boldsymbol{v}_3$.（这个组合中的权重和为 1.）当 c_2,c_3 和 t 满足 $(1-c_2-c_3)\boldsymbol{v}_1+c_2\boldsymbol{v}_2+c_3\boldsymbol{v}_3=\boldsymbol{a}+t\boldsymbol{b}$ 时，射线 $\boldsymbol{x}(t)$ 与平面相交.

重排得到 $c_2(\boldsymbol{v}_2-\boldsymbol{v}_1)+c_3(\boldsymbol{v}_3-\boldsymbol{v}_1)+t(-\boldsymbol{b})=\boldsymbol{a}-\boldsymbol{v}_1$. 矩阵形式为：

$$\begin{bmatrix}\boldsymbol{v}_2-\boldsymbol{v}_1 & \boldsymbol{v}_3-\boldsymbol{v}_1 & -\boldsymbol{b}\end{bmatrix}\begin{bmatrix}c_2\\c_3\\t\end{bmatrix}=\boldsymbol{a}-\boldsymbol{v}_1$$

对这里给出的具体点

$$\boldsymbol{v}_2-\boldsymbol{v}_1=\begin{bmatrix}7\\0\\2\end{bmatrix},\quad \boldsymbol{v}_3-\boldsymbol{v}_1=\begin{bmatrix}4\\10\\4\end{bmatrix},\quad \boldsymbol{a}-\boldsymbol{v}_1=\begin{bmatrix}-1\\-1\\16\end{bmatrix}$$

做增广矩阵行简化运算

$$\begin{bmatrix}7&4&-0.7&-1\\0&10&-0.4&-1\\2&4&3&16\end{bmatrix}\sim\begin{bmatrix}1&0&0&0.3\\0&1&0&0.1\\0&0&1&0.5\end{bmatrix}$$

因此，$c_2=0.3,c_3=0.1$ 和 $t=5$. 从而，相交点是

$$\boldsymbol{x}(5)=\boldsymbol{a}+5\boldsymbol{b}=\begin{bmatrix}0\\0\\10\end{bmatrix}+5\begin{bmatrix}0.7\\0.4\\-3\end{bmatrix}=\begin{bmatrix}3.5\\2.0\\-5.0\end{bmatrix}$$

㊀ 参见 Joshua Fender and Jonathan Rose，"A High-Speed Ray Tracing Engine Built on a Field-Programmable System"，in *Proc. Int.Conf on Field-Programmable Technology*，IEEE（2003）.（一个单处理器每秒可以计算 60 亿次射线-三角相交.）

同样，

$$x(5)=(1-0.3-0.1)v_1+0.3v_2+0.1v_3$$
$$=0.6\begin{bmatrix}1\\1\\-6\end{bmatrix}+0.3\begin{bmatrix}8\\1\\-4\end{bmatrix}+0.1\begin{bmatrix}5\\11\\-2\end{bmatrix}=\begin{bmatrix}3.5\\2.0\\-5.0\end{bmatrix}$$

由于 $x(5)$ 的重心权重都是正的，所以相交点位于三角形内部. ■

练习题

1. 试用一种快速的方法来确定什么时候三个点是不共线的.
2. 点 $v_1=\begin{bmatrix}4\\1\end{bmatrix}, v_2=\begin{bmatrix}1\\0\end{bmatrix}, v_3=\begin{bmatrix}5\\4\end{bmatrix}$ 和 $v_4=\begin{bmatrix}1\\2\end{bmatrix}$ 构成了一个仿射相关集，求出权重 $c_1,\cdots,c_4$，构造出一个**仿射相关关系** $c_1v_1+\cdots+c_4v_4=\mathbf{0}$，这里 $c_1+\cdots+c_4=0$，并且 c_i 不全为零.（提示：见定理 5 证明的结尾.）

习题 8.2

在习题 1~6 中，确定这些点的集合是否为仿射相关的. 如果是，构造这些点处的一仿射相关关系.

1. $\begin{bmatrix}3\\-3\end{bmatrix},\begin{bmatrix}0\\6\end{bmatrix},\begin{bmatrix}2\\0\end{bmatrix}$

2. $\begin{bmatrix}2\\1\end{bmatrix},\begin{bmatrix}5\\4\end{bmatrix},\begin{bmatrix}-3\\-2\end{bmatrix}$

3. $\begin{bmatrix}1\\2\\-1\end{bmatrix},\begin{bmatrix}-2\\-4\\8\end{bmatrix},\begin{bmatrix}2\\-1\\11\end{bmatrix},\begin{bmatrix}0\\15\\-9\end{bmatrix}$

4. $\begin{bmatrix}-2\\5\\3\end{bmatrix},\begin{bmatrix}0\\-3\\7\end{bmatrix},\begin{bmatrix}1\\-2\\-6\end{bmatrix},\begin{bmatrix}-2\\7\\-3\end{bmatrix}$

5. $\begin{bmatrix}1\\0\\-2\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix},\begin{bmatrix}-1\\5\\1\end{bmatrix},\begin{bmatrix}0\\5\\-3\end{bmatrix}$

6. $\begin{bmatrix}1\\3\\1\end{bmatrix},\begin{bmatrix}0\\-1\\-2\end{bmatrix},\begin{bmatrix}2\\5\\2\end{bmatrix},\begin{bmatrix}3\\5\\0\end{bmatrix}$

在习题 7 和 8 中，求出 p 的相对于前面点的仿射无关集的重心坐标.

7. $\begin{bmatrix}1\\-1\\2\\1\end{bmatrix},\begin{bmatrix}2\\1\\0\\1\end{bmatrix},\begin{bmatrix}1\\2\\-2\\0\end{bmatrix}, p=\begin{bmatrix}5\\4\\-2\\2\end{bmatrix}$

8. $\begin{bmatrix}0\\1\\-2\\1\end{bmatrix},\begin{bmatrix}1\\1\\0\\2\end{bmatrix},\begin{bmatrix}1\\4\\-6\\5\end{bmatrix}, p=\begin{bmatrix}-1\\1\\-4\\0\end{bmatrix}$

在习题 9 和 10 中，判断每一个命题的真假，并说明理由.

9. a. 如果 $v_1,\cdots,v_p$ 在 $\mathbb{R}^n$ 中，$\{v_1-v_2,v_3-v_2,\cdots,v_p-v_2\}$ 是线性相关的，那么 $\{v_1,\cdots,v_p\}$ 是仿射相关的.
   b. 如果 $v_1,\cdots,v_p$ 在 $\mathbb{R}^n$ 中，集合的齐次形式 $\{\tilde{v}_1,\cdots,\tilde{v}_p\}$ 在 $\mathbb{R}^{n+1}$ 中是线性无关的，那么 $\{v_1,\cdots,v_p\}$ 是仿射相关的.
   c. 如果存在不全为零的实数 $c_1,\cdots,c_k$，使 $c_1+\cdots+c_k=1$ 且 $c_1v_1+\cdots+c_kv_k=\mathbf{0}$，那么点集 $\{v_1,\cdots,v_k\}$ 是仿射相关的.
   d. 如果 $S=\{v_1,\cdots,v_p\}$ 是 $\mathbb{R}^n$ 中的一个仿射无关集，$\mathbb{R}^n$ 中 p 有由 S 确定的负的重心坐标，那么 p 不在 aff S 中.
   e. 如果 v_1,v_2,v_3,a 和 b 在 $\mathbb{R}^3$ 中，对 $t\geqslant 0$，$a+tb$ 与顶点为 v_1,v_2 和 v_3 的三角形相交，那么相交点的重心坐标全都是非负的.
10. a. 如果 $\{v_1,\cdots,v_p\}$ 是 $\mathbb{R}^n$ 中的一个仿射相关集，那么其齐次形式 $\{\tilde{v}_1,\cdots,\tilde{v}_p\}$ 在 $\mathbb{R}^{n+1}$ 中可能是线性无关的.
    b. 如果 v_1,v_2,v_3 和 v_4 在 $\mathbb{R}^3$ 中，集合 $\{v_2-v_1,v_3-v_1,v_4-v_1\}$ 是线性无关的，那么 $\{v_1,\cdots,v_4\}$ 是仿射无关的.
    c. 给定 $\mathbb{R}^n$ 中集合 $S=\{b_1,\cdots,b_k\}$，aff S 中每个点 p 都有一个 $b_1,\cdots,b_k$ 的仿射组合的唯一表达.

d. 如果 $\mathbb{R}^3$ 中一个三角形的每一个顶点 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 都给出了详细的颜色信息，那么可以运用 $\boldsymbol{p}$ 的重心坐标给出 $\text{aff}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 中插入点 $\boldsymbol{p}$ 的颜色.

e. 如果 $\mathbb{R}^2$ 中 T 是一个三角形，$\boldsymbol{p}$ 位于三角形的边缘，则 $\boldsymbol{p}$ 的重心坐标不全是正的.

11. 解释为什么 $\mathbb{R}^3$ 中 5 个或更多的点构成的集合必是仿射相关的.

12. 证明：当 $p \geqslant n+2$ 时，$\mathbb{R}^n$ 中一个集合 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 是仿射相关的.

13. 只用仿射相关的定义来证明：当且仅当 $\boldsymbol{v}_1=\boldsymbol{v}_2$ 时，$\mathbb{R}^n$ 中一个指标集 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是仿射相关的.

14. 仿射相关的条件比线性相关的条件更强，因此，一个仿射相关集自动是线性相关的. 同样，一个线性无关集一定不是仿射相关的，从而必是仿射无关的. 构造 $\mathbb{R}^2$ 中两个线性相关的指标集 S_1 和 S_2，使得 S_1 是仿射相关的，S_2 是仿射无关的. 在每一种情况下，集合或者包含 1 个、2 个或 3 个非零点.

15. 令 $\boldsymbol{v}_1=\begin{bmatrix}-1\\2\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}0\\4\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}2\\0\end{bmatrix}$，$S=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$.

a. 证明：集合 S 是仿射无关的.

b. 求点 $\boldsymbol{p}_1=\begin{bmatrix}2\\3\end{bmatrix},\boldsymbol{p}_2=\begin{bmatrix}1\\2\end{bmatrix},\boldsymbol{p}_3=\begin{bmatrix}-2\\1\end{bmatrix}$，$\boldsymbol{p}_4=\begin{bmatrix}1\\-1\end{bmatrix}$ 和 $\boldsymbol{p}_5=\begin{bmatrix}1\\1\end{bmatrix}$ 关于 S 的重心坐标.

c. 令 T 是含有顶点 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 和 $\boldsymbol{v}_3$ 的三角形. 当把 T 的边延长后，直线把 $\mathbb{R}^2$ 分成 7 个区域，见图 8-13. 注意每一个区域中点的重心坐标的符号. 例如，$\boldsymbol{p}_5$ 在三角形 T 中，它的所有重心坐标是正的. 点 $\boldsymbol{p}_5$ 有坐标 $(-,+,+)$. 由于 $\boldsymbol{p}_1$ 在穿过 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的直线的 $\boldsymbol{v}_3$ 侧，所以它的第三坐标是正的. 由于 $\boldsymbol{p}_1$ 在穿过 $\boldsymbol{v}_2$ 和 $\boldsymbol{v}_3$ 的直线的 $\boldsymbol{v}_1$ 反侧，它的第一坐标是负的. 点 $\boldsymbol{p}_2$ 在 T 的 $\boldsymbol{v}_2\boldsymbol{v}_3$ 边缘上，它的坐标是 $(0,+,+)$. 不需要计算实际值，确定图 8-13 所示的点 $\boldsymbol{p}_6$，$\boldsymbol{p}_7$ 和 $\boldsymbol{p}_8$ 的重心坐标的符号.

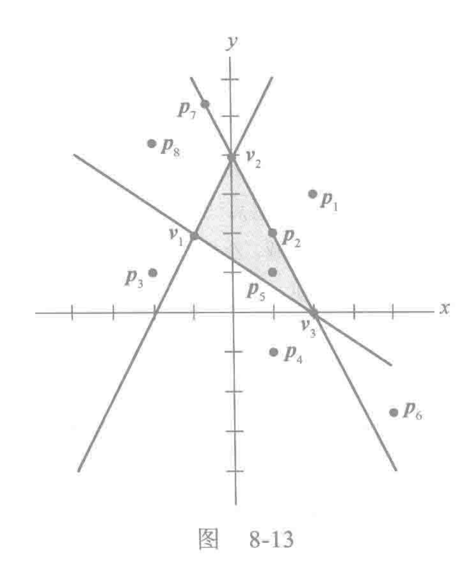


图 8-13

16. 令 $\boldsymbol{v}_1=\begin{bmatrix}0\\1\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}1\\5\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}4\\3\end{bmatrix},\boldsymbol{p}_1=\begin{bmatrix}3\\5\end{bmatrix}$，$\boldsymbol{p}_2=\begin{bmatrix}5\\1\end{bmatrix}$，$\boldsymbol{p}_3=\begin{bmatrix}2\\3\end{bmatrix},\boldsymbol{p}_4=\begin{bmatrix}-1\\0\end{bmatrix},\boldsymbol{p}_5=\begin{bmatrix}0\\4\end{bmatrix}$，$\boldsymbol{p}_6=\begin{bmatrix}1\\2\end{bmatrix},\boldsymbol{p}_7=\begin{bmatrix}6\\4\end{bmatrix}$，$S=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$.

a. 证明：集合 S 是仿射无关的.

b. 求 $\boldsymbol{p}_1,\boldsymbol{p}_2$ 和 $\boldsymbol{p}_3$ 关于 S 的重心坐标.

c. 画出顶点为 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 和 $\boldsymbol{v}_3$ 的三角形 T，如图 8-10 一样延长它的边，并画出点 $\boldsymbol{p}_4,\boldsymbol{p}_5,\boldsymbol{p}_6$ 和 $\boldsymbol{p}_7$，不需要计算出实际的值，确定点 $\boldsymbol{p}_4,\boldsymbol{p}_5,\boldsymbol{p}_6$ 和 $\boldsymbol{p}_7$ 的重心坐标的符号.

17. 对 $\mathbb{R}^n$ 中的一个仿射无关集合 $S=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k\}$，证明定理 6.（提示：一种方法是模仿 4.4 节中的定理 7 的证明.）

18. 令 T 是位于“标准”位置的四面体，三条边沿着 $\mathbb{R}^3$ 中的三个正向坐标轴，并假定顶点是 $a\boldsymbol{e}_1,b\boldsymbol{e}_2,c\boldsymbol{e}_3$ 和 $\boldsymbol{0}$，这里 $[\boldsymbol{e}_1\quad\boldsymbol{e}_2\quad\boldsymbol{e}_3]=I_3$. 求 $\mathbb{R}^3$ 中任意一点 $\boldsymbol{p}$ 的重心坐标的公式.

19. 令 $\{\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3\}$ 是 $\mathbb{R}^n$ 中仿射相关集合中的一

点，并令 $f:\mathbb{R}^n \to \mathbb{R}^m$ 是一个线性变换．证明：$\{f(\boldsymbol{p}_1), f(\boldsymbol{p}_2), f(\boldsymbol{p}_3)\}$ 是 $\mathbb{R}^m$ 中仿射相关的点集．

20. 假定 $\{\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3\}$ 是 $\mathbb{R}^n$ 中一个仿射无关集，$\boldsymbol{q}$ 是 $\mathbb{R}^n$ 中的任意点．证明：平移集 $\{\boldsymbol{p}_1+\boldsymbol{q}, \boldsymbol{p}_2+\boldsymbol{q}, \boldsymbol{p}_3+\boldsymbol{q}\}$ 也是仿射无关的．

在习题 21~24 中，$\boldsymbol{a}, \boldsymbol{b}$ 和 $\boldsymbol{c}$ 是 $\mathbb{R}^2$ 中不共线的点，$\boldsymbol{p}$ 是 $\mathbb{R}^2$ 中任意其他点．令 $\triangle \boldsymbol{abc}$ 表示由 $\boldsymbol{a}, \boldsymbol{b}$ 和 $\boldsymbol{c}$ 确定的三角形的区域，$\triangle \boldsymbol{pbc}$ 表示由 $\boldsymbol{p}$，$\boldsymbol{b}$ 和 $\boldsymbol{c}$ 确定的区域．为了方便，假定 $\boldsymbol{a}, \boldsymbol{b}$ 和 $\boldsymbol{c}$ 被重新安排使得 $\det[\tilde{\boldsymbol{a}}\ \tilde{\boldsymbol{b}}\ \tilde{\boldsymbol{c}}]$ 是正的，这里 $\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{b}}$ 和 $\tilde{\boldsymbol{c}}$ 是这些点的标准齐次形式．

21. 证明：$\triangle \boldsymbol{abc}$ 的面积就是 $\det[\tilde{\boldsymbol{a}}\quad \tilde{\boldsymbol{b}}\quad \tilde{\boldsymbol{c}}]/2$．（提示：查阅 3.2 节和 3.3 节，包括习题．）

22. 令 $\boldsymbol{p}$ 是穿过 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 的直线上的点，证明：$\det[\tilde{\boldsymbol{a}}\quad \tilde{\boldsymbol{b}}\quad \tilde{\boldsymbol{p}}]=0$

23. 令 $\boldsymbol{p}$ 是 $\triangle \boldsymbol{abc}$ 内的任意点，重心坐标为 (r,s,t)，从而

$$[\tilde{\boldsymbol{a}}\quad \tilde{\boldsymbol{b}}\quad \tilde{\boldsymbol{c}}]\begin{bmatrix} r \\ s \\ t \end{bmatrix} = \tilde{\boldsymbol{p}}$$

运用习题 19 和关于行列式的一个事实（第 3 章）来证明：

$r = (\triangle \boldsymbol{pbc}\text{的面积})/(\triangle \boldsymbol{abc}\text{的面积})$

$s = (\triangle \boldsymbol{apc}\text{的面积})/(\triangle \boldsymbol{abc}\text{的面积})$

$t = (\triangle \boldsymbol{abp}\text{的面积})/(\triangle \boldsymbol{abc}\text{的面积})$

24. 取 $\boldsymbol{q}$ 是从 $\boldsymbol{b}$ 到 $\boldsymbol{c}$ 的直线段上的点，并考虑穿过 $\boldsymbol{q}$ 和 $\boldsymbol{a}$ 的直线，它也可被写成 $\boldsymbol{p}=(1-x)\boldsymbol{q}+x\boldsymbol{a}$，$x$ 为实数．证明：对每一个 x，$\det[\tilde{\boldsymbol{p}}\quad \tilde{\boldsymbol{b}}\quad \tilde{\boldsymbol{c}}] = x\cdot\det[\tilde{\boldsymbol{a}}\quad \tilde{\boldsymbol{b}}\quad \tilde{\boldsymbol{c}}]$．从这个结论和已有的结论可得出，参数 x 是 $\boldsymbol{p}$ 的第一个重心坐标．然而，由前面的构造可知，参数 x 也确定了 $\boldsymbol{p}$ 和 $\boldsymbol{q}$ 之间的沿着从 $\boldsymbol{q}$ 到 $\boldsymbol{a}$ 的直线段的距离关系．（当 x=1 时，$\boldsymbol{p}$=$\boldsymbol{a}$．）当把这个事实运用到例 5 时，表明顶点 $\boldsymbol{a}$ 和点 $\boldsymbol{q}$ 的颜色可以确定一个光滑地插入沿着从 $\boldsymbol{a}$ 到 $\boldsymbol{q}$ 的直线运动的 $\boldsymbol{p}$ 点的颜色．

练习题答案

1. 从例 1 可知，问题就是确定这些点是否是仿射相关．运用例 2 中的方法，从两个点中减去另一点，如果所得到的两个新点中的一个是另一个的倍数，那么原来的三个点在同一条直线上，否则不共线．

2. 定理 5 的证明表明，点之间的仿射相关关系对应于这些点的齐次形式的线性相关关系，所用的权重值相同．因此，行简化得：

$$\begin{bmatrix} \tilde{\boldsymbol{v}}_1 & \tilde{\boldsymbol{v}}_2 & \tilde{\boldsymbol{v}}_3 & \tilde{\boldsymbol{v}}_4 \end{bmatrix} = \begin{bmatrix} 4 & 1 & 5 & 1 \\ 1 & 0 & 4 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 4 & 1 & 5 & 1 \\ 1 & 0 & 4 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1.25 \\ 0 & 0 & 1 & 0.75 \end{bmatrix}$$

把所得矩阵看作是含有 4 个变量的 $A\boldsymbol{x}=\boldsymbol{0}$ 的系数矩阵，则 x_4 是自由未知量，$x_1 = x_4, x_2 = -1.25x_4$ 和 $x_3 = -0.75x_4$．一个解是 $x_1 = x_4 = 4, x_2 = -5$ 和 $x_3 = -3$．齐次形式之间的线性相关关系就是 $4\tilde{\boldsymbol{v}}_1 - 5\tilde{\boldsymbol{v}}_2 - 3\tilde{\boldsymbol{v}}_3 + 4\tilde{\boldsymbol{v}}_4 = \boldsymbol{0}$．从而 $4\boldsymbol{v}_1 - 5\boldsymbol{v}_2 - 3\boldsymbol{v}_3 + 4\boldsymbol{v}_4 = \boldsymbol{0}$．

另一个求解的方法就是把问题平移到原点处来处理，为此从其他点中都减去 $\boldsymbol{v}_1$，求平移所得的这些新点之间的一个线性相关关系，然后重排表达式．这种方法的运算量与上面所用的方法大致相同．

8.3 凸组合

8.1 节中考察了一个特殊的线性组合形式

$$c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_k\boldsymbol{v}_k\text{，其中 }c_1+c_2+\cdots+c_k=1$$

本节将限制这些权重为非负.

定义 $\mathbb{R}^n$ 中点 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_k$ 的一个凸组合就是如下形式的线性组合:

$$c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_k\boldsymbol{v}_k$$

对于所有的 i，有 $c_1+c_2+\cdots+c_k=1$ 和 $c_i\geqslant 0$. 一个集合 S 中所有凸组合的集称为 S 的**凸包**，记为 conv S.

单点 $\boldsymbol{v}_1$ 的凸包就是集合 $\{\boldsymbol{v}_1\}$，与仿射包相同. 在其他的情形下，凸包真包含在仿射包中，回忆不同点 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 之间的仿射包就是直线

$$\boldsymbol{y}=(1-t)\boldsymbol{v}_1+t\boldsymbol{v}_2,\quad t\in\mathbb{R}$$

由于一个凸组合的权重是非负的，$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 中的点可写为

$$\boldsymbol{y}=(1-t)\boldsymbol{v}_1+t\boldsymbol{v}_2,\quad 0\leqslant t\leqslant 1$$

这是在点 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 之间的**线段**，记为 $\overline{\boldsymbol{v}_1\boldsymbol{v}_2}$.

如果一个集合 S 是仿射无关的，且 $\boldsymbol{p}\in\text{aff }S$，那么当且仅当 $\boldsymbol{p}$ 的重心坐标是非负时，$\boldsymbol{p}\in\text{conv }S$. 例 1 证明了一个特殊的情形，其中 S 是仿射无关的.

例 1 令

$$\boldsymbol{v}_1=\begin{bmatrix}3\\0\\6\\-3\end{bmatrix},\quad \boldsymbol{v}_2=\begin{bmatrix}-6\\3\\3\\0\end{bmatrix},\quad \boldsymbol{v}_3=\begin{bmatrix}3\\6\\0\\3\end{bmatrix},\quad \boldsymbol{p}_1=\begin{bmatrix}0\\3\\3\\0\end{bmatrix},\quad \boldsymbol{p}_2=\begin{bmatrix}-10\\5\\11\\-4\end{bmatrix}$$

且 $S=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$. 注意 S 是正交集. 确定 $\boldsymbol{p}_1$ 是否在 Span S，aff S 及 conv S 中. 同样考虑 $\boldsymbol{p}_2$ 的情况.

解 若 $\boldsymbol{p}_1$ 至少是 S 上点的线性组合，由于 S 是正交集合，则很容易求出权重. 令 W 为 S 生成的子空间. 类似于 6.3 节的计算证明了 $\boldsymbol{p}_1$ 在 W 上的正交投影 $\boldsymbol{p}_1$ 是本身:

$$\begin{aligned}\text{proj}_W\,\boldsymbol{p}_1&=\frac{\boldsymbol{p}_1\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1}\boldsymbol{v}_1+\frac{\boldsymbol{p}_1\cdot\boldsymbol{v}_2}{\boldsymbol{v}_2\cdot\boldsymbol{v}_2}\boldsymbol{v}_2+\frac{\boldsymbol{p}_3\cdot\boldsymbol{v}_3}{\boldsymbol{v}_3\cdot\boldsymbol{v}_3}\boldsymbol{v}_3\\&=\frac{18}{54}\boldsymbol{v}_1+\frac{18}{54}\boldsymbol{v}_2+\frac{18}{54}\boldsymbol{v}_3\\&=\frac{1}{3}\begin{bmatrix}3\\0\\6\\-3\end{bmatrix}+\frac{1}{3}\begin{bmatrix}-6\\3\\3\\0\end{bmatrix}+\frac{1}{3}\begin{bmatrix}3\\0\\6\\3\end{bmatrix}=\frac{1}{3}\begin{bmatrix}0\\3\\3\\0\end{bmatrix}=\boldsymbol{p}_1\end{aligned}$$

由此证明 $\boldsymbol{p}_1$ 在 Span S 内. 同时由于系数和为 1, $\boldsymbol{p}_1$ 在 S 的仿射包内. 进一步, 由于系数是非负的, $\boldsymbol{p}_1$ 在 S 的凸包内.

对于 $\boldsymbol{p}_2$. 类似的计算证明 $\text{proj}_W \boldsymbol{p}_2 \neq \boldsymbol{p}_2$. 由于 $\text{proj}_W \boldsymbol{p}_2$ 是 Span S 上到 $\boldsymbol{p}_2$ 最近的点, 所以, 点 $\boldsymbol{p}_2$ 不在 Span S 上. 特别地, $\boldsymbol{p}_2$ 不可能在 S 的仿射包内或凸包内. ■

回忆一下, 集合 S 是仿射的, 若 S 内任意两点确定的直线都在 S 里. 而在凸集里, 要考虑的是线段而不是直线.

定义 集合 S 是凸的, 若对于 S 上的每个 $\boldsymbol{p},\boldsymbol{q} \in S$, 线段 $\overline{\boldsymbol{pq}}$ 在 S 里.

直觉上, 若集合 S 里的每两点都可以相互"看到", 且视线不越出该集合, 那么集合 S 是凸的. 图 8-14 说明了这个思想.

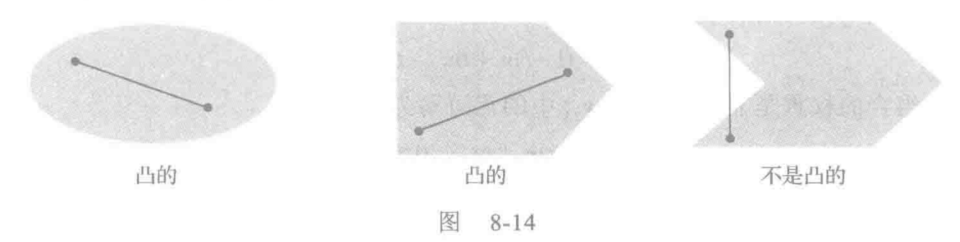


图 8-14

下面的结论类似于仿射集的定理 2.

定理 7 集合 S 是凸集当且仅当 S 中的点的凸组合在 S 中, 即 S 是凸集当且仅当 $S = \text{conv}\, S$.

证明 讨论类似于定理 2 的证明, 这里只给出有所不同的证明部分. 取 $k+1$ 个点的凸组合, 考虑 $\boldsymbol{y} = c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k + c_{k+1}\boldsymbol{v}_{k+1}$, 这里 $c_1 + c_2 + \cdots + c_{k+1} = 1$, 且对所有的 i, $0 \leqslant c_i \leqslant 1$. 若 $c_{k+1} = 1$, 则 $\boldsymbol{y} = \boldsymbol{v}_{k+1}$ 属于 S, 不需要进一步证明. 若 $c_{k+1} < 1$, 令 $t = c_1 + \cdots + c_k$, 则 $t = 1 - c_{k+1} > 0$ 且

$$\boldsymbol{y} = (1 - c_{k+1})\left(\frac{c_1}{t}\boldsymbol{v}_1 + \cdots + \frac{c_k}{t}\boldsymbol{v}_k\right) + c_{k+1}\boldsymbol{v}_{k+1} \tag{1}$$

由归纳假设, 点 $\boldsymbol{z}=(c_1/t)\boldsymbol{v}_1+\cdots+(c_k/t)\boldsymbol{v}_k$ 在 S 内, 由于非负系数和为 1. 于是, 方程 (1) 表明 $\boldsymbol{y}$ 为 S 内两点的凸组合. 由归纳法原理, S 内的点的每个凸组合在 S 内. ■

下面的定理 9 提供了集合凸包的进一步的几何特征. 它需要有关交集的初步结论. 回顾 4.1 节 (习题 32) 中的一个事实——两个子空间的交集本身仍是一个子空间. 事实上, 任一组子空间的交集本身仍是一个子空间. 对于仿射集与凸集有相似的结论.

定理 8 设 $\{S_\alpha : \alpha \in \mathcal{A}\}$ 是任一组凸集, 则 $\cap_{\alpha\in\mathcal{A}} S_\alpha$ 是凸集; 若 $\{T_\beta : \beta \in \mathcal{B}\}$ 是任一组仿射集, 则 $\cap_{\beta\in\mathcal{B}} S_\beta$ 是仿射集.

证明 若 $\boldsymbol{p},\boldsymbol{q} \in \cap S_\alpha$, 则 $\boldsymbol{p},\boldsymbol{q}$ 在每个 S_α 中. 由于每个 S_α 是凸集, 连接 $\boldsymbol{p},\boldsymbol{q}$ 的线段在每个 S_α 中, 因此线段属于 $\cap S_\alpha$. 对于仿射集的证明与此类似.

定理 9 对任何集合 S, S 的凸包是所有包含 S 的凸集的交集.

证明 令 T 为包含 S 的所有凸集的交集. 由于 S 的凸包是包含 S 的一个凸集，则 $T\subset \text{conv}\, S$. 另一方面，令 C 是包含 S 的任意一个凸集，则 C 包含 C 中所有点构成的凸组合（定理 7），因此也包含子集 S 中所有点构成的凸组合，即 $\text{conv}\, S\subset C$. 由于对包含 S 的每个凸集 C 都成立，故对所有包含 S 的凸集的交集同样成立，即 $\text{conv}\, S\subset T$. ■

定理 9 证明了 S 的凸包是包含 S 的“最小”的凸集，这一点诠释了“凸包”的自然含义. 例如，考虑平面 $\mathbb{R}^2$ 中可被某个大长方形包含的点集 S，想象环绕 S 外边沿拉伸橡皮筋带. 由于橡皮筋限制于 S，它构架了 S 凸组合的边界.或者用另一个模拟，S 的凸组合填满 S 内部所有的洞及 S 边界所有的凹陷.

例 2

a. $\mathbb{R}^2$ 中集合 S，T 的凸包如下图所示.

b. 令 S 是三维空间 $\mathbb{R}^3$ 的标准基构成的集合，$S=\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}$. 则凸集 S 是 $\mathbb{R}^2$ 中顶点为 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$ 的一个三角面. 见图 8-15.

例 3 令 $S=\left\{\begin{bmatrix}x\\y\end{bmatrix}: x\geqslant 0, y=x^2\right\}$，证明 S 的凸包是原点和 $\left\{\begin{bmatrix}x\\y\end{bmatrix}: x>0, y\geqslant x^2\right\}$ 的并集. 如图 8-16 所示.

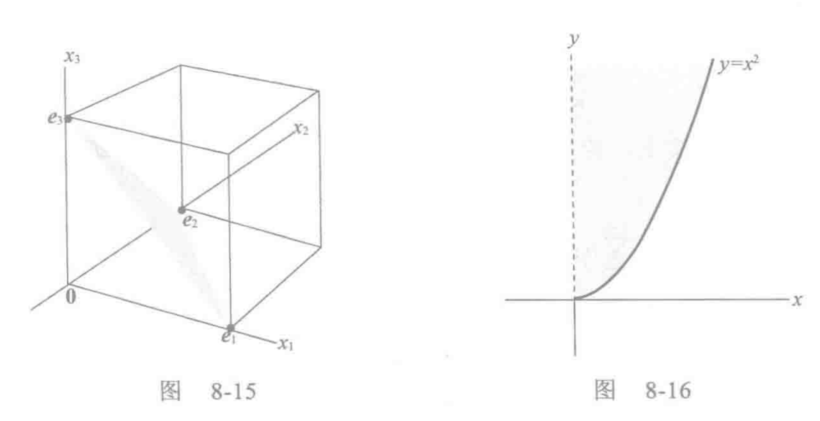


图 8-15 图 8-16

解 S 的凸包中每点都在连接 S 中两点的线段上. 图 8-16 中的虚线表明，除了原点外，y 的正半轴不在 S 的凸包上，原点是 S 在 y 轴上唯一的点. 图 8-16 所示的 S 的凸包似乎是合理的，但是如何确定点 $(10^{-2},10^4)$ 在连接原点与 S 中另一点的线段上？

考虑图 8-16 阴影部分的任意点 $\boldsymbol{p}$，令

$$\boldsymbol{p}=\begin{bmatrix}a\\b\end{bmatrix},\quad a>0 \text{ 且 } b\geqslant a^2$$

经过 **0** 与 $\boldsymbol{p}$ 的直线满足方程 $y=(b/a)t$，t 为实数. 直线与 S 的交点满足$(b/a)t=t^2$，即 $t=b/a$. 因此$\boldsymbol{p}$在连接 **0** 与$\begin{bmatrix} b/a \\ b^2/a^2 \end{bmatrix}$的线段上. 由此证明图 8-16 正确. ■

下面的定理是凸集学习中的基本定理，于 1907 年首次由 Constantin Caratheodory 证明. 若 $\boldsymbol{p}$ 在 S 的凸包中，则按定义，$\boldsymbol{p}$ 是 S 的点的一个凸组合. 但定义没有规定需要 S 中多少个点构成这个组合. Caratheodory 的著名定理表明在 n 维空间中，凸组合中 S 的点的个数不必多于 $n+1$.

定理 10（Caratheodory） *若 S 是 $\mathbb{R}^n$ 中一非空子集，则 S 的凸包中的每一点可以由 S 中 $n+1$ 个或更少的点的凸组合表示.*

证明 假定 $\boldsymbol{p}$ 在 S 的凸包中，则 $\boldsymbol{p}=c_1\boldsymbol{v}_1+\cdots+c_k\boldsymbol{v}_k$，这里 $\boldsymbol{v}_i\in S$，$c_1+\cdots+c_k=1$，且 $c_i\geqslant 0, i=1,\cdots,k$. 我们的目的就是证明 $\boldsymbol{p}$ 存在这样的表达式且 $k\leqslant n+1$.

若 $k>n+1$，则由 8.2 节习题 12 知 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k\}$ 是仿射相关的，因此存在不全为 0 的标量 $d_1,\cdots,d_k$，使得

$$\sum_{i=1}^{k} d_i\boldsymbol{v}_i=\boldsymbol{0}\text{且}\sum_{i=1}^{k} d_i=0$$

考虑如下两个方程：

$$c_1\boldsymbol{v}_1+\cdots+c_k\boldsymbol{v}_k=\boldsymbol{p}$$
$$d_1\boldsymbol{v}_1+\cdots+d_k\boldsymbol{v}_k=\boldsymbol{0}$$

第一个方程减去第二个方程的适当倍数，我们可以消去一个 $\boldsymbol{v}_i$，从而得到 S 少于 k 个点构成的凸组合，即该凸组合等价于 $\boldsymbol{p}$.

由于并非所有的系数 d_i 为 0. 我们可以假定（若需要重新排下标）$d_k>0$ 且对那些 $d_i>0$ 的所有 i 有 $c_k/d_k\leqslant c_i/d_i$，对 $i=1$，…，k，令 $b_i=c_i-(c_k/d_k)d_i$. 则 $b_k=0$ 且

$$\sum_{i=1}^{k} b_i=\sum_{i=1}^{k} c_i-\frac{c_k}{d_k}\sum_{i=1}^{k} d_i=1-0=1$$

进一步，每个 $b_i\geqslant 0$. 事实上，若 $d_i\leqslant 0$，则 $b_i\geqslant c_i\geqslant 0$. 若 $d_i>0$，则 $b_i=d_i(c_i/d_i-c_k/d_k)\geqslant 0$. 由构造，得

$$\sum_{i=1}^{k-1} b_i\boldsymbol{v}_i=\sum_{i=1}^{k} b_i\boldsymbol{v}_i=\sum_{i=1}^{k}\left(c_i=\frac{c_k}{d_k}d_i\right)\boldsymbol{v}_i$$
$$=\sum_{i=1}^{k} c_i\boldsymbol{v}_i-\frac{c_k}{d_k}\sum_{i=1}^{k} d_i\boldsymbol{v}_i=\sum_{i=1}^{k} c_i\boldsymbol{v}_i=\boldsymbol{p}$$

因此$\boldsymbol{p}$是 $k-1$ 个点 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_{k-1}$ 的凸组合. 重复上述步骤到$\boldsymbol{p}$是由 S 中 $n+1$ 个点构成的凸组合. ■

下面的例子说明了上述证明中的计算.

例 4 令

$$\boldsymbol{v}_1=\begin{bmatrix}1\\0\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}2\\3\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}5\\4\end{bmatrix},\boldsymbol{v}_4=\begin{bmatrix}3\\0\end{bmatrix},\boldsymbol{p}=\begin{bmatrix}\frac{10}{3}\\\frac{5}{2}\end{bmatrix}$$

且 $S=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4\}$. 则

$$\frac{1}{4}\boldsymbol{v}_1+\frac{1}{6}\boldsymbol{v}_2+\frac{1}{2}\boldsymbol{v}_3+\frac{1}{12}\boldsymbol{v}_4=\boldsymbol{p} \tag{2}$$

使用 Caratheodory 定理的证明过程把 $\boldsymbol{p}$ 表示为 S 中三个点构成的凸组合.

解 集合 S 是仿射相关的. 使用 8.2 节中的方法得到一个仿射相关的关系式.

$$-5\boldsymbol{v}_1+4\boldsymbol{v}_2-3\boldsymbol{v}_3+4\boldsymbol{v}_4=\boldsymbol{0} \tag{3}$$

接下来，选择（3）式中的点 $\boldsymbol{v}_2,\boldsymbol{v}_4$，它们的系数为正. 对上述点，计算方程（2）与（3）系数的比例，$\boldsymbol{v}_2$ 的系数比为 $\frac{1}{6}\div 4=\frac{1}{24}$，而 $\boldsymbol{v}_4$ 的为 $\frac{1}{12}\div 4=\frac{1}{48}$. $\boldsymbol{v}_4$ 的比更小，消去 $\boldsymbol{v}_4$：

$$\left(\frac{1}{4}+\frac{5}{48}\right)\boldsymbol{v}_1+\left(\frac{1}{6}-\frac{4}{48}\right)\boldsymbol{v}_2+\left(\frac{1}{2}+\frac{3}{48}\right)\boldsymbol{v}_3+\left(\frac{1}{12}-\frac{4}{48}\right)\boldsymbol{v}_4=\boldsymbol{p}$$

$$\frac{17}{48}\boldsymbol{v}_1+\frac{4}{48}\boldsymbol{v}_2+\frac{27}{48}\boldsymbol{v}_3=\boldsymbol{p}$$

■

一般而言，通过减少需要的点的个数并不能改善这个结果. 事实上，$\mathbb{R}^2$ 中任意三个非共线的点构成的三角形的质心在这三个点的凸包中，而不是在任意两点的凸包中.

练习题

1. 令 $\boldsymbol{v}_1=\begin{bmatrix}6\\2\\2\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}7\\1\\5\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}-2\\4\\-1\end{bmatrix},\boldsymbol{p}_1=\begin{bmatrix}1\\3\\1\end{bmatrix}$，及 $\boldsymbol{p}_2=\begin{bmatrix}3\\2\\1\end{bmatrix}$，令 $S=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$. 确定 $\boldsymbol{p}_1$ 和 $\boldsymbol{p}_2$ 是否在 S 的凸集中.
2. 令 S 为曲线 $y=1/x(x>0)$ 上的点集，用几何特征解释 S 的凸包由曲线 S 上及 S 的上方的点构成.

习题 8.3

1. 在 $\mathbb{R}^2$ 中，令 $S=\left\{\begin{bmatrix}0\\y\end{bmatrix}:0\leqslant y<1\right\}\cup\left\{\begin{bmatrix}2\\0\end{bmatrix}\right\}$. 描述 S 的凸包.
2. 描述 $\mathbb{R}^2$ 中点 $\begin{bmatrix}x\\y\end{bmatrix}$ 构成的集合 S 满足如下给定条件的凸包. 证明你的结论.（证明 S 中任意一点 $\boldsymbol{p}$ 属于 S 的凸包.）
   a. $y=1/x$ 且 $x\geqslant 1/2$
   b. $y=\sin x$
   c. $y=x^{1/2}$ 且 $x\geqslant 0$
3. 考虑 8.1 节习题 5 中的点. $\boldsymbol{p}_1$，$\boldsymbol{p}_2$ 和 $\boldsymbol{p}_3$ 哪个在 S 的凸包中？
4. 考虑 8.1 节习题 6 中的点. $\boldsymbol{p}_1$，$\boldsymbol{p}_2$ 和 $\boldsymbol{p}_3$ 哪个在 S 的凸包中？
5. 令 $\boldsymbol{v}_1=\begin{bmatrix}-1\\-3\\4\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}0\\-3\\1\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}1\\-1\\4\end{bmatrix},\boldsymbol{v}_4=\begin{bmatrix}1\\1\\-2\end{bmatrix}$，$\boldsymbol{p}_1=\begin{bmatrix}1\\-1\\2\end{bmatrix}$，$\boldsymbol{p}_2=\begin{bmatrix}0\\-2\\2\end{bmatrix}$，且 $S=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4\}$. 确定 $\boldsymbol{p}_1$，$\boldsymbol{p}_2$ 是否在 S 的凸包中.
6. 令 $\boldsymbol{v}_1=\begin{bmatrix}2\\0\\-1\\2\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}0\\-2\\2\\1\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}-2\\1\\0\\2\end{bmatrix},\boldsymbol{p}_1=\begin{bmatrix}-1\\2\\-\frac{3}{2}\\\frac{5}{2}\end{bmatrix}$，$\boldsymbol{p}_2=\begin{bmatrix}-\frac{1}{2}\\0\\\frac{1}{4}\\\frac{7}{4}\end{bmatrix},\boldsymbol{p}_3=\begin{bmatrix}6\\-4\\1\\-1\end{bmatrix},\boldsymbol{p}_4=\begin{bmatrix}-1\\-2\\0\\4\end{bmatrix}$，且 $S=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 为正交集. 确定每一 $\boldsymbol{p}_i$ 是否属于 Span S，aff S 或 conv S.
   a. $\boldsymbol{p}_1$ b. $\boldsymbol{p}_2$ c. $\boldsymbol{p}_3$ d. $\boldsymbol{p}_4$

习题 7~10 使用 8.2 节中的术语.

7. a. 令 $T=\left\{\begin{bmatrix}-1\\0\end{bmatrix},\begin{bmatrix}2\\3\end{bmatrix},\begin{bmatrix}4\\1\end{bmatrix}\right\}$，令 $\boldsymbol{p}_1=\begin{bmatrix}2\\1\end{bmatrix},\boldsymbol{p}_2=\begin{bmatrix}3\\2\end{bmatrix}$，

$p_3=\begin{bmatrix}2\\0\end{bmatrix}$, $p_4=\begin{bmatrix}0\\2\end{bmatrix}$. 求出 p_1，p_2，p_3，p_4 相对于 T 的重心坐标.

b. 由（a）的答案确定 p_1，p_2，p_3，p_4 在 T 的凸包（即三角形）内部、外部还是边上？

8. 重复习题 7 的过程，这里 $T=\left\{\begin{bmatrix}2\\0\end{bmatrix},\begin{bmatrix}0\\5\end{bmatrix},\begin{bmatrix}-1\\1\end{bmatrix}\right\}$，

令 $p_1=\begin{bmatrix}2\\1\end{bmatrix}, p_2=\begin{bmatrix}1\\1\end{bmatrix}, p_3=\begin{bmatrix}1\\ \frac{1}{3}\end{bmatrix}, p_4=\begin{bmatrix}1\\0\end{bmatrix}$.

9. 令 $S=\{v_1,v_2,v_3,v_4\}$ 是仿射不相关集，考虑点 $p_1,\cdots,p_5$ 相对于 S 的重心坐标分别为 $(2,0,0,-1)$, $\left(0,\frac{1}{2},\frac{1}{4},\frac{1}{4}\right)$, $\left(\frac{1}{2},0,\frac{3}{2},-1\right)$, $\left(\frac{1}{3},\frac{1}{4},\frac{1}{4},\frac{1}{6}\right)$ 和 $\left(\frac{1}{3},0,\frac{2}{3},0\right)$. 确定每点 $p_1,\cdots,p_5$ 是在凸包 S 一个四面体内部、外部，还是表面. 这 5 个点中存在点在凸包 S 的边上吗？

10. 重复习题 9 的过程，这里 $q_1,\cdots,q_5$ 相对于 S 的质心坐标分别为 $\left(\frac{1}{8},\frac{1}{4},\frac{1}{8},\frac{1}{2}\right)$，$\left(\frac{3}{4},-\frac{1}{4},0,\frac{1}{2}\right)$，$\left(0,\frac{3}{4},\frac{1}{4},0\right)$，$(0,-2,0,3)$ 和 $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3},0\right)$.

判断习题 11 和 12 中每个命题的真假，并说明理由.

11. a. 若 $y=c_1v_1+c_2v_2+c_3v_3$，且 $c_1+c_2+c_3=1$，则 y 是 v_1,v_2,v_3 的凸组合.

b. 若 S 是非空集合，则 S 的凸包包含一些不属于 S 的点.

c. 若 S，T 都是凸集，则 $S\cup T$ 也是凸集.

12. a. 集合 S 是凸集，若 $x,y\in S$，则 x 和 y 间的线段在 S 内.

b. 若 S，T 都是凸集，则 $S\cap T$ 也是凸集.

c. 若 S 是 $\mathbb{R}^5$ 中的非空集合，且 $y\in \text{conv}S$，则存在 S 中不同的点 $v_1,\cdots,v_6$, 使得 y 是 $v_1,\cdots,v_6$ 的凸组合.

13. 令 S 是 $\mathbb{R}^n$ 中的凸子集,假定 $f:\mathbb{R}^n\to\mathbb{R}^m$ 是一线性变换. 证明：集合 $f(S)=\{f(x):x\in S\}$ 是 $\mathbb{R}^m$ 中的凸子集.

14. 令 $f:\mathbb{R}^n\to\mathbb{R}^m$ 是一线性变换，且 T 是 $\mathbb{R}^m$ 中的凸子集. 证明：集合 $S=\{x\in\mathbb{R}^n: f(x)\in T\}$ 是 $\mathbb{R}^n$ 的凸子集.

15. 令 $v_1=\begin{bmatrix}1\\0\end{bmatrix}, v_2=\begin{bmatrix}1\\2\end{bmatrix}, v_3=\begin{bmatrix}4\\2\end{bmatrix}, v_4=\begin{bmatrix}4\\0\end{bmatrix}, p=\begin{bmatrix}2\\1\end{bmatrix}$. 证实 $p=\frac{1}{3}v_1+\frac{1}{3}v_2+\frac{1}{6}v_3+\frac{1}{6}v_4$ 且 $v_1-v_2+v_3-v_4=\mathbf{0}$. 使用 Caratheodory 定理的证明过程把 p 表示为 v_i 中三个点构成的凸组合. 使用两种方法.

16. 重复习题 9 的过程，点 $v_1=\begin{bmatrix}-1\\0\end{bmatrix}, v_2=\begin{bmatrix}0\\3\end{bmatrix}, v_3=\begin{bmatrix}3\\1\end{bmatrix}, v_4=\begin{bmatrix}1\\-1\end{bmatrix}, p=\begin{bmatrix}1\\2\end{bmatrix}$.

已知 $p=\frac{1}{121}v_1+\frac{72}{121}v_2+\frac{37}{121}v_3+\frac{1}{11}v_4$ 和 $10v_1-6v_2+7v_3-11v_4=\mathbf{0}$.

在习题 17~20 中，证明关于 $\mathbb{R}^n$ 中子集 A 与 B 的命题. 证明中可能会使用前面习题中的结论.

17. 若 $A\subset B$ 且 B 是凸集，则 $\text{conv}\,A\subset B$.

18. 若 $A\subset B$，则 $\text{conv}\,A\subset\text{conv}\,B$.

19. a. [（conv A）∪（conv B）]⊂conv（$A\cup B$）.

b. 在 $\mathbb{R}^2$ 找一个例子，以证明（a）中等号未必成立.

20. a. conv（$A\cap B$）⊂[（conv A）∩（conv B）].

b. 在 $\mathbb{R}^2$ 找一个例子，以证明（a）中等号未必成立.

21. 令 p_0,p_1,p_2 是 $\mathbb{R}^n$ 中的点，且定义 $f_0(t)=(1-t)p_0+tp_1$，$f_1(t)=(1-t)p_1+tp_2$，$g(t)=(1-t)f_0(t)+tf_1(t)$，$0\leqslant t\leqslant 1$. 各点如图所示，画出说明 $f_0\left(\frac{1}{2}\right)$，$f_1\left(\frac{1}{2}\right)$ 和 $g\left(\frac{1}{2}\right)$ 的图.

22. 重复习题 21 的过程，画出说明 $f_0\left(\frac{3}{4}\right)$，$f_1\left(\frac{3}{4}\right)$ 和 $g\left(\frac{3}{4}\right)$ 的图.

23. 令 $g(t)$ 为习题 21 中的定义，它的图像称为二

次贝塞尔曲线，常运用于某些计算机图形设计. p_0, p_1, p_2 称为曲线的控制点，计算仅包含 p_0, p_1, p_2 的公式 $g(t)$. 那么 $g(t)$ 在 conv$\{p_0, p_1, p_2\}$ 中，$0 \leqslant t \leqslant 1$.

24. 给定 $\mathbb{R}^n$ 中的控制点 p_0, p_1, p_2, p_3，令 $g_1(t)$ 为习题 23 中 p_0, p_1, p_2 确定的二次贝塞尔曲线，$0 \leqslant t \leqslant 1$，令 $g_2(t)$ 为 p_1, p_2, p_3 确定的二次贝塞尔曲线. 定义 $h(t) = (1-t)g_1(t) + tg_2(t)$，$0 \leqslant t \leqslant 1$. 证明 $h(t)$ 的图形在这四个控制点的凸包中. 这条曲线称为三次贝塞尔曲线，它的定义基于构建贝塞尔曲线的算法.（在稍后 8.6 节讨论.）k 维的贝塞尔曲线由 $k+1$ 个控制点确定，它的图形位于这些控制点的凸包中.

练习题答案

1. 点 v_1, v_2, v_3 非正交，计算

$$v_2 - v_1 = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}, v_3 - v_1 = \begin{bmatrix} -8 \\ 2 \\ -3 \end{bmatrix}, p_1 - v_1 = \begin{bmatrix} -5 \\ 1 \\ -1 \end{bmatrix}, p_2 - v_1 = \begin{bmatrix} -3 \\ 0 \\ -1 \end{bmatrix}$$

把 $p_1 - v_1$ 和 $p_2 - v_1$ 加到矩阵 $[v_2 - v_1 \quad v_3 - v_1]$ 中作行变换：

$$\begin{bmatrix} 1 & -8 & -5 & -3 \\ -1 & 2 & 1 & 0 \\ 3 & -3 & -1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & \frac{1}{3} & 1 \\ 0 & 1 & \frac{2}{3} & \frac{1}{2} \\ 0 & 0 & 0 & -\frac{5}{2} \end{bmatrix}$$

第三列说明 $p_1 - v_1 = \frac{1}{3}(v_2 - v_1) + \frac{2}{3}(v_3 - v_1)$，推出 $p_1 = 0v_1 + \frac{1}{3}v_2 + \frac{2}{3}v_3$. 故 p_1 在 S 的凸包中. 事实上，p_1 在点集 $\{v_2, v_3\}$ 的凸包中.

矩阵的最后一列说明 $p_2 - v_1$ 不是 $v_2 - v_1$ 与 $v_3 - v_1$ 的线性组合，所以 p_2 不是 v_1, v_2, v_3 的仿射组合，故 p_2 不在 S 的凸包中.

另一解法就是把齐次形式的增广矩阵行化简为阶梯矩阵：

$$[\tilde{v}_1 \quad \tilde{v}_2 \quad \tilde{v}_3 \quad \tilde{p}_1 \quad \tilde{p}_2] \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 & \frac{2}{3} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

2. 若 p 是曲线 S 上方一点，则穿过点 p 斜率为-1 的直线在到达 x,y 的正半轴前与 S 交于两点.

8.4 超平面

超平面在 $\mathbb{R}^n$ 空间中有着特别的作用，由于它们把空间分成两个不相交的部分，正如一个平面把 $\mathbb{R}^3$ 分成两部分以及一条直线切开 $\mathbb{R}^2$. 探讨超平面的一个关键要领是对直线和平面启用简单的隐式表达，而不是像之前讨论仿射集时所用的显式，即参数形式表达.[㊀]

$\mathbb{R}^2$ 中直线的隐式方程为 $ax + by = d$，$\mathbb{R}^3$ 中平面的隐式方程为 $ax + by + cz = d$. 两方程把直线

㊀ 1.5 节介绍了参数形式表达.

或平面都描述为一个线性表示式（也叫线性函数）取得固定值 d 时所有点的集合.

定义 $\mathbb{R}^n$ 上的一个线性函数是从 $\mathbb{R}^n$ 到 $\mathbb{R}$ 的一个线性变换 f, 对 $\mathbb{R}$ 上的每个标量 d, 符号 $[f{:}d]$ 表示 $\mathbb{R}^n$ 中使得 f 的值为 d 的所有点的集合，即

$$[f{:}d] \text{是集合} \{x \in \mathbb{R}^n : f(x) = d\}$$

零函数是对 $\mathbb{R}^n$ 上所有点 $\boldsymbol{x}$ 都有 $f(\boldsymbol{x}) = 0$ 的线性函数. $\mathbb{R}^n$ 上所有其他的线性函数都称为**非零函数**.

例 1 在 $\mathbb{R}^2$ 上，直线 $x - 4y = 13$ 是 $\mathbb{R}^2$ 上的超平面. 它是所有使得函数 $f(x, y) = x - 4y$ 的值等于 13 的点组成的集合，即直线是集合 $[f{:}13]$.

例 2 在 $\mathbb{R}^3$ 上，平面 $5x - 2y + 3z = 21$ 是 $\mathbb{R}^3$ 上的一个超平面，它是所有使得函数 $g(x, y, z) = 5x - 2y + 3z$ 的值等于 21 的点组成的集合，即超平面是集合 $[g{:}21]$.

若 f 是 $\mathbb{R}^n$ 上的线性函数，那么这个线性变换 f 的标准矩阵是一个 $1 \times n$ 矩阵 A. 令 $A = [a_1\ a_2 \cdots a_n]$，则

$$[f{:}0] \text{等同于} \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = 0\} = \text{Nul}\ A \tag{1}$$

若 f 是非零函数，由秩的定理[⊖]，则 A 的秩为 1，$\dim \text{NUl}\ A = n-1$，因此 $[f{:}0]$ 是 $n-1$ 维的，从而是一个超平面. 同时，若 d 是 $\mathbb{R}$ 中任意数，则

$$[f{:}d] \text{等同于} \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = d\} \tag{2}$$

回顾 1.5 节中的定理 6，$A\boldsymbol{x} = \boldsymbol{b}$ 的解集是将 $A\boldsymbol{x} = \boldsymbol{0}$ 的解集沿 $A\boldsymbol{x} = \boldsymbol{b}$ 的一特解 $\boldsymbol{p}$ 平移得到的. 当 A 为变换 f 的标准矩阵时，定理表明

$$[f{:}d] = [f{:}0] + \boldsymbol{p}, \text{对} [f{:}d] \text{中任意的} \boldsymbol{p} \tag{3}$$

故集合 $[f{:}d]$ 是平行于 $[f{:}0]$ 的超平面. 见图 8-17.

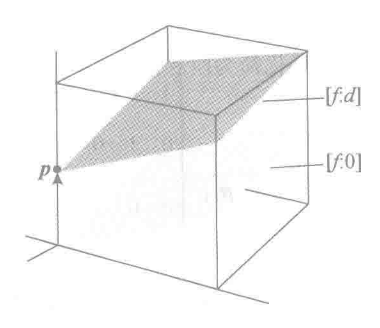


图 8-17 平行超平面，$f(\boldsymbol{p})=d$

当 A 为 $1 \times n$ 矩阵，方程 $A\boldsymbol{x} = d$ 的左边可写成内积 $\boldsymbol{n} \cdot \boldsymbol{x}$，$\boldsymbol{n} \in \mathbb{R}^n$ 且它的元素与 A 的相同，故由（2），

$$[f{:}d] \text{等同于} \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{n} \cdot \boldsymbol{x} = d\} \tag{4}$$

于是，$[f:0] = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{n} \cdot \boldsymbol{x} = 0\}$，此式表明 $[f:0]$ 是由 $\boldsymbol{n}$ 生成的子空间的正交补. 用 $\mathbb{R}^3$ 中微积分

⊖ 参见 2.9 节的定理 14 或 4.6 节的定理 14.

和几何术语，$\boldsymbol{n}$称为$[f:0]$的**法向量**.（这里“法向量”的含义并不要求它是单位向量.）同时$\boldsymbol{n}$正交于每个平行的超平面$[f:d]$，即使当$d\neq 0$时$\boldsymbol{n}\cdot\boldsymbol{x}$不为零.

$[f:d]$的另一个名称为f的水平集，当对所有的$\boldsymbol{x}$有$f(\boldsymbol{x})=\boldsymbol{n}\cdot\boldsymbol{x}$时，$\boldsymbol{n}$有时称为$f$的梯度.

例 3 令$\boldsymbol{n}=\begin{bmatrix}3\\4\end{bmatrix},\boldsymbol{v}=\begin{bmatrix}1\\-6\end{bmatrix},H=\{\boldsymbol{x}:\boldsymbol{n}\cdot\boldsymbol{x}=12\}$，故$H=[f:12]$，这里$f(x,y)=3x+4y$. 故$H$为直线$3x+4y=12$. 求平行的超平面（直线）$H_1=H+\boldsymbol{v}$的隐式表达式.

解 首先，求出H_1中的一点$\boldsymbol{p}$. 取H中一点，把它加上$\boldsymbol{v}$. 例如H中点$\begin{bmatrix}0\\3\end{bmatrix}$，故$\boldsymbol{p}=\begin{bmatrix}1\\-6\end{bmatrix}+\begin{bmatrix}0\\3\end{bmatrix}=\begin{bmatrix}1\\-3\end{bmatrix}$在$H_1$中. 现在计算$\boldsymbol{n}\cdot\boldsymbol{p}=-9$，这表明$H_1=[f:-9]$，见图 8-18，该图也表明子空间$H_0=\{\boldsymbol{x}:\boldsymbol{n}\cdot\boldsymbol{x}=0\}$.

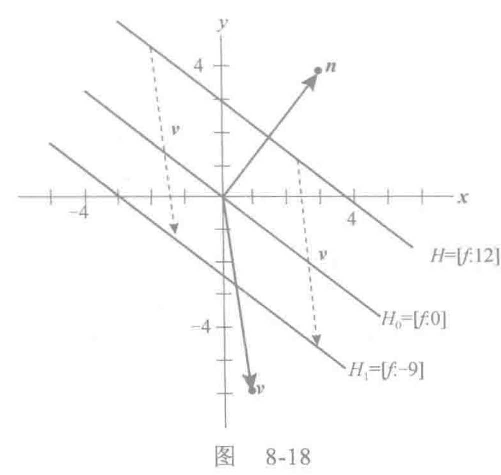


图 8-18

接下来三个例子说明了超平面中隐式与显式表达式的联系. 例 4 以一个隐式开始.

例 4 在$\mathbb{R}^2$上，以参数向量的形式给出直线$x-4y=13$的显式表达式.

解 这等价于解方程$A\boldsymbol{x}=\boldsymbol{b}$，这里$A=[1\quad -4]$，且$b$的值为 13. 改写等式$x=13+4y$，这里$y$是一个自由变量. 用参数形式，解为

$$\boldsymbol{x}=\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}13+4y\\y\end{bmatrix}=\begin{bmatrix}13\\0\end{bmatrix}+y\begin{bmatrix}4\\1\end{bmatrix}=\boldsymbol{p}+y\boldsymbol{q},\ y\in\mathbb{R}.$$ ■

反过来，把直线的显式转换成隐式的过程较复杂一些. 其基本思想是构造$[f:0]$，然后求出$[f:d]$的d.

例 5 令$\boldsymbol{v}_1=\begin{bmatrix}1\\2\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}6\\0\end{bmatrix}$，且令$L_1$是经过$\boldsymbol{v}_1$，$\boldsymbol{v}_2$的直线. 求线性函数$f$及常数$d$，使得$L_1=[f:d]$.

解 直线L_1平行于经过点$\boldsymbol{v}_2-\boldsymbol{v}_1$及原点的直线$L_0$. 定义$L_0$的方程如下：

$$[a\quad b]\begin{bmatrix}x\\y\end{bmatrix}=0\text{ 或 }\boldsymbol{n}\cdot\boldsymbol{x}=0,\text{ 这里 }\boldsymbol{n}=\begin{bmatrix}a\\b\end{bmatrix}\qquad(5)$$

由于 $\boldsymbol{n}$ 与包含点 $\boldsymbol{v}_2-\boldsymbol{v}_1$ 的子空间 L_0 正交，计算

$$\boldsymbol{v}_2-\boldsymbol{v}_1=\begin{bmatrix}6\\0\end{bmatrix}-\begin{bmatrix}1\\2\end{bmatrix}=\begin{bmatrix}5\\-2\end{bmatrix}$$

并解

$$[a\ \ b]\begin{bmatrix}5\\-2\end{bmatrix}=0$$

通过观察，得一个解为$[a\quad b]=[2\quad 5]$. 令 $f(x,y)=2x+5y$. 由（5），$L_0=[f{:}0]$且对某个 d 有 $L_1=[f{:}d]$. 由于 $\boldsymbol{v}_1$ 在直线 L_1 上，$d=f(\boldsymbol{v}_1)=2(1)+5(2)=12$. 故 L_1 的方程为 $2x+5y=12$. 作为检验，注意到 $f(\boldsymbol{v}_2)=f(6,0)=2(6)+5(0)=12$. 所以 $\boldsymbol{v}_2$ 也在直线 L_1 上. ■

例 6 令 $\boldsymbol{v}_1=\begin{bmatrix}1\\1\\1\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}2\\-1\\4\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}3\\1\\2\end{bmatrix}$, 求过点 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 的平面 H_1 的隐式表达式$[f{:}d]$.

解 H_1 平行于平面 H_0，H_0 过原点及平移点

$$\boldsymbol{v}_2-\boldsymbol{v}_1=\begin{bmatrix}1\\-2\\3\end{bmatrix}及\boldsymbol{v}_3-\boldsymbol{v}_1=\begin{bmatrix}2\\0\\1\end{bmatrix}$$

由于这两点线性无关，$H_0=\operatorname{Span}\{\boldsymbol{v}_2-\boldsymbol{v}_1,\ \boldsymbol{v}_3-\boldsymbol{v}_1\}$. 令 $\boldsymbol{n}=\begin{bmatrix}a\\b\\c\end{bmatrix}$ 与 H_0 正交. 那么 $\boldsymbol{v}_2-\boldsymbol{v}_1$，$\boldsymbol{v}_3-\boldsymbol{v}_1$ 分别与 $\boldsymbol{n}$ 正交. 也就是说，$(\boldsymbol{v}_2-\boldsymbol{v}_1)\cdot\boldsymbol{n}=0$，$(\boldsymbol{v}_3-\boldsymbol{v}_1)\cdot\boldsymbol{n}=0$. 这两个方程构成的方程组的增广矩阵可进行行变换：

$$[1\ \ -2\ \ 3]\begin{bmatrix}a\\b\\c\end{bmatrix}=0,[2\ \ 0\ \ 1]\begin{bmatrix}a\\b\\c\end{bmatrix}=0,\begin{bmatrix}1&-2&3&0\\2&0&1&0\end{bmatrix}$$

进行行操作得到 $a=(-\frac{2}{4})c$，$b=(-\frac{5}{4})c$，c是自由变量. 如取 $c=4$. 那么 $\boldsymbol{n}=\begin{bmatrix}-2\\5\\4\end{bmatrix}$ 且 $H_0=[f:0]$，这里 $f(\boldsymbol{x})=-2x_1+5x_2+4x_3$.

平行超平面 H_1 就是$[f:d]$. 为了求 d，由 $\boldsymbol{v}_1\in H_1$，计算 $d=f(\boldsymbol{v}_1)=f(1,1,1)=-2(1)+5(1)+4(1)=7$. 作为检验，计算

$$f(\boldsymbol{v}_2)=f(2,-1,4)=-2(2)+5(-1)+4(4)=7\cdot$$ ■

例 6 的过程可推广到高维空间，然而对于特别情形 $\mathbb{R}^3$ 也可以使用**叉积公式**计算 $\boldsymbol{n}$，并使用行列式记号作为辅助记忆工具：

$$\begin{aligned}\boldsymbol{n}&=(\boldsymbol{v}_2-\boldsymbol{v}_1)\times(\boldsymbol{v}_3-\boldsymbol{v}_1)\\&=\begin{vmatrix}1&2&\boldsymbol{i}\\-2&0&\boldsymbol{j}\\3&1&\boldsymbol{k}\end{vmatrix}=\begin{vmatrix}-2&0\\3&1\end{vmatrix}\boldsymbol{i}-\begin{vmatrix}1&2\\3&1\end{vmatrix}\boldsymbol{j}+\begin{vmatrix}1&2\\-2&0\end{vmatrix}\boldsymbol{k}\end{aligned}$$

$$= -2\boldsymbol{i} + 5\boldsymbol{j} + 4\boldsymbol{k} = \begin{bmatrix} -2 \\ 5 \\ 4 \end{bmatrix}$$

若只需要f的计算公式，叉积的计算写成一般的行列式形式：

$$f(x_1, x_2, x_3) = \begin{vmatrix} 1 & 2 & x_1 \\ -2 & 0 & x_2 \\ 3 & 1 & x_3 \end{vmatrix} = \begin{vmatrix} -2 & 0 \\ 3 & 1 \end{vmatrix} x_1 - \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} x_2 + \begin{vmatrix} 1 & 2 \\ -2 & 0 \end{vmatrix} x_3 = -2x_1 + 5x_2 + 4x_3$$

目前为止，被考察的每个超平面可描述成$[f:d]$，f 为某个线性函数，d 为 $\mathbb{R}$ 中某个数，或等价地描述为$\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{n} \cdot \boldsymbol{x} = d\}$，$\boldsymbol{n}$ 为 $\mathbb{R}$ 中某个数. 下面的定理说明了每个超平面都有这些等价描述.

定理 11 $\mathbb{R}^n$ 中的子集 H 是超平面当且仅当 $H = [f:d]$，f 为某个非零线性函数，d 为 $\mathbb{R}$ 中某个数. 因而若 H 是超平面，则存在非零向量 $\boldsymbol{n}$ 与实数 d，使得 $H = \{\boldsymbol{x} : \boldsymbol{n} \cdot \boldsymbol{x} = d\}$.

证明 假定 H 是超平面，取 $\boldsymbol{p} \in H$，且令 $H_0 = H - \boldsymbol{p}$. 则 H_0 是一个 $n-1$ 维的子空间. 接下来取 $\boldsymbol{y} \notin H_0$，由 6.3 节中的正交分解定理，

$$\boldsymbol{y} = \boldsymbol{y}_1 + \boldsymbol{n}$$

这里 $\boldsymbol{y}_1$ 是 H_0 中一个向量，$\boldsymbol{n}$ 与 H_0 中每个向量正交. 函数f定义为

$$f(\boldsymbol{x}) = \boldsymbol{n} \cdot \boldsymbol{x},\ \ \boldsymbol{x} \in \mathbb{R}^n$$

由内积性质可知，这是一个线性函数. 由 $\boldsymbol{n}$ 的含义可知，$[f:0]$是包含 H_0 的超平面. 从而，

$$H_0 = [f:0]$$

[讨论：H_0 包含一个由 $n-1$ 个向量构成的基 S，而由于 S 在 $n-1$ 维子空间$[f:0]$中，根据基定理可知，S 也是$[f:0]$的一个基.]最后，令 $d = f(\boldsymbol{p}) = \boldsymbol{n} \cdot \boldsymbol{p}$，则如前面的（3）式所示，

$$[f:d] = [f:0] + \boldsymbol{p} = H_0 + \boldsymbol{p} = H$$

反过来，若f是一线性函数，d是一实数，则由上面的（1）和（3）可知，$[f:d]$是一个超平面.

■

超平面的许多重要应用依赖于超平面能够“分割”两个集合. 直观上，这意味着其中一个集合在超平面的一面，而另一集合在另一面. 下面的术语将更清晰地表达这一思想.

> $\mathbb{R}^n$ 中的拓扑：术语与事实
>
> 对 $\mathbb{R}^n$ 中任意点 $\boldsymbol{p}$，及任意实数 $\delta > 0$，以 $\boldsymbol{p}$ 为心 δ 为半径的**开球** $B(\boldsymbol{p}, \delta)$ 表示为
>
> $$B(\boldsymbol{p}, \delta) = \{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{p}\| < \delta\}$$
>
> 设 S 为 $\mathbb{R}^n$ 中的集合，点 $\boldsymbol{p}$ 为 S 的**内点**，若存在 $\delta > 0$，使得 $B(\boldsymbol{p}, \delta) \subset S$. 若以 $\boldsymbol{p}$ 为心的每个开球既与 S 相交又与 S 的补集相交，则 $\boldsymbol{p}$ 称为 S 的**边界点**. 若集合不包含边界点，则集合称为**开集**.（这等价于说 S 中的所有点都为 S 的内点.）若集合包含所有边界点，则称它为**闭集**.（若 S 包含部分但不是所有边界点，则 S 既非开集也非闭集.）若存在 $\delta > 0$ 使得 $S \subset B(\boldsymbol{0}, \delta)$，则集合 S 称为**有界的**. 若 $\mathbb{R}^n$ 中的集合既是闭的又是有界的，则称为**紧致的**.

> **定理** 开集的凸包是开集，紧致集的凸包是紧致的.（闭集的凸包不一定是闭的，见习题 27.）

例 7 令 $S=\text{conv}\left\{\begin{bmatrix}-2\\2\end{bmatrix},\begin{bmatrix}-2\\-2\end{bmatrix},\begin{bmatrix}2\\-2\end{bmatrix},\begin{bmatrix}2\\2\end{bmatrix}\right\}$，$\boldsymbol{p}_1=\begin{bmatrix}-1\\0\end{bmatrix}$，及 $\boldsymbol{p}_2=\begin{bmatrix}2\\1\end{bmatrix}$，如图 8-19 所示. 则 $\boldsymbol{p}_1$ 是内点，这是由于 $B(\boldsymbol{p},\frac{3}{4})\subset S$. 点 $\boldsymbol{p}_2$ 是边界点，这是由于以 $\boldsymbol{p}_2$ 为心的每一开球与 S 和 S 的补集都相交. 由于 S 包含所有边界点，故 S 是闭集. 由于 $S\subset B(\boldsymbol{0},3)$，所以集合 S 是有界的. 因此 S 也是紧致的. ■

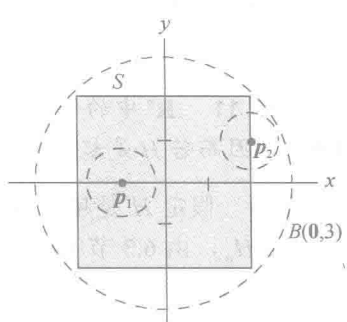


图 8-19 集合 S 是闭的和有界的

注释：若 f 是线性函数，则 $f(A)\leqslant d$ 指的是 $f(\boldsymbol{x})\leqslant d, \boldsymbol{x}\in A$. 对于不等式是反向的或者严格的情况，也有相应的注释.

定义 超平面 $H=[f:d]$ 若满足下列条件之一：

(i) $f(A)\leqslant d$ 且 $f(B)\geqslant d$

(ii) $f(A)\geqslant d$ 且 $f(B)\leqslant d$

则该超平面被**分割**成两个集合 A 与 B. 若在以上条件中，所有的弱不等式变为严格不等式，则称 H **严格分割**集合 A 与 B.

注意，严格分割需要两集合不相交，而分割不需要这个条件. 事实上，若平面上的两圆是外切的，则它们共同的切线分割它们（但不是严格分割）.

为了严格分割两集合，需要两集合不相交，即使对于闭凸集，此条件也是不充分的. 例如，令

$$A=\left\{\begin{bmatrix}x\\y\end{bmatrix}:x\geqslant\frac{1}{2},\frac{1}{x}\leqslant y\leqslant 2\right\} \text{ 且 } B=\left\{\begin{bmatrix}x\\y\end{bmatrix}:x\geqslant 0,y=0\right\}$$

那么 A 与 B 是不相交的闭凸集，但它们不能由超平面（$\mathbb{R}^2$ 中直线）严格分割. 见图 8-20. 超平面分割（或严格分割）两集合的问题比它刚出现时要复杂.

集合 A 与 B 上有许多有趣的条件，蕴含超平面的存在性，而如下两个定理对本节来说已足够. 第一个定理的证明需要相当多的基础知识[1]，而第二个定理可由第一个简单地推出.

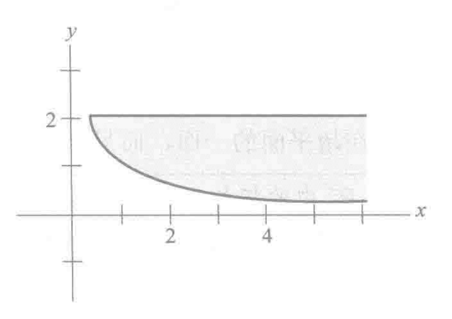


图 8-20 不相交闭凸集

定理 12 设 A 与 B 是非空凸集，且 A 是紧致的，B 是闭的. 那么存在超平面 H 严格分割 A 与 B，当且仅当 $A\cap B=\varnothing$.

[1] 定理 12 的证明见 Steven R. Lay，*Convex Sets and Their Applications*(New York:John Wiley&Sons,1982;Mineola,NY: Dover Publications,2007),pp.34-39.

定理 13 设 A 与 B 是非空紧致集. 那么存在超平面严格分割 A 与 B，当且仅当 $(\text{conv } A)\cap(\text{conv } B)=\varnothing$.

证明 设 $(\text{conv } A)\cap(\text{conv } B)=\varnothing$. 由于紧致集的凸包是紧致的，定理 12 保证了存在超平面 H 严格分割 conv A 与 conv B. 显然，H 严格分割更小的集合 A 与 B.

反之，假定超平面 $H=[f:d]$ 严格分割 A 与 B. 不失一般性，设 $f(A)<d$ 且 $f(B)>d$，令 $\boldsymbol{x}=c_1\boldsymbol{x}_1+\cdots+c_k\boldsymbol{x}_k$ 是 A 元素的任一凸组合，则

$$f(\boldsymbol{x})=c_1 f(\boldsymbol{x}_1)+\cdots+c_k f(\boldsymbol{x}_k)<c_1 d+\cdots c_k d=d$$

这是由于 $c_1+\cdots+c_k=1$. 从而 $f(\text{conv } A)<d$. 同样 $f(\text{conv } B)>d$，故 $H=[f:d]$ 严格分割 conv A 与 conv B. 由定理 12，conv A 与 conv B 是不相交的. ■

例 8 设 $\boldsymbol{a}_1=\begin{bmatrix}2\\1\\1\end{bmatrix},\boldsymbol{a}_2=\begin{bmatrix}-3\\2\\1\end{bmatrix},\boldsymbol{a}_3=\begin{bmatrix}3\\4\\0\end{bmatrix},\boldsymbol{b}_1=\begin{bmatrix}1\\0\\2\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}2\\-1\\5\end{bmatrix}$，且设 $A\{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3\}$，$B=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$. 证明超平面 $H=[f:5]$ 不能分割 A 与 B，这里 $f(x_1,x_2,x_3)=2x_1-3x_2\ \ +x_3$. 存在平行于 H 的超平面分割 A 与 B 吗？A 与 B 的凸包相交吗？

解 计算线性函数 f 在集合 A 与 B 中每点的值：

$f(\boldsymbol{a}_1)<3$，$f(\boldsymbol{a}_2)=-11$，$f(\boldsymbol{a}_3)=-6$，$f(\boldsymbol{b}_1)=4$，$f(\boldsymbol{b}_2)=12$

由于 $f(\boldsymbol{b}_1)=4$ 比 5 小，$f(\boldsymbol{b}_2)=12$ 比 5 大，B 中的点位于集合 $H=[f:5]$ 的两侧，故 H 不分割 A 与 B.

由于 $f(A)<3$ 且 $f(B)>3$，平行超平面 $[f:3]$ 严格分割 A 与 B. 由定理 13，$(\text{conv } A)\cap(\text{conv } B)=\varnothing$.

注意：若不存在平行于 H 的超平面严格分割 A 与 B，这并非蕴含它们的凸包相交，可能存在不平行于 H 的超平面严格分割 A 与 B.

练习题

设 $\boldsymbol{p}_1=\begin{bmatrix}1\\0\\2\end{bmatrix},\boldsymbol{p}_2=\begin{bmatrix}-1\\2\\1\end{bmatrix},\boldsymbol{n}_1=\begin{bmatrix}1\\1\\-2\end{bmatrix},\boldsymbol{n}_2=\begin{bmatrix}-2\\1\\3\end{bmatrix}$，且设 H_1 是 $\mathbb{R}^3$ 中的超平面，过点 $\boldsymbol{p}_1$ 且与 $\boldsymbol{n}_1$ 正交，H_2 是过点 $\boldsymbol{p}_2$ 且与 $\boldsymbol{n}_2$ 正交的超平面. 求 $H_1\cap H_2$ 的显式表达式以说明如何产生 $H_1\cap H_2$ 中的所有点.

习题 8.4

1. 设 L 是 $\mathbb{R}^2$ 中过点 $\begin{bmatrix}-1\\4\end{bmatrix}$ 和 $\begin{bmatrix}3\\1\end{bmatrix}$ 的一条直线. 求线性函数 f 及实数 d 满足 $L=[f:d]$.
2. 设 L 是 $\mathbb{R}^2$ 中过点 $\begin{bmatrix}1\\4\end{bmatrix}$ 和 $\begin{bmatrix}-2\\-1\end{bmatrix}$ 的一条直线. 求线性函数 f 及实数 d 满足 $L=[f:d]$.

在习题 3 和 4 中确定每个集合是开集还是闭集或既非开集也非闭集.

3. a. $\{(x,y):y>0\}$
   b. $\{(x,y):x=2,1\leqslant y\leqslant 3\}$
   c. $\{(x,y):x=2,1<y<3\}$
   d. $\{(x,y):xy=1,x>0\}$
   e. $\{(x,y):xy\geqslant 1,x>0\}$

4. a. $\{(x,y): x^2+y^2=1\}$

   b. $\{(x,y): x^2+y^2>1\}$

   c. $\{(x,y): x^2+y^2\leqslant 1, y>0\}$

   d. $\{(x,y): y\geqslant x^2\}$

   e. $\{(x,y): y<x^2\}$

在习题 5 和 6 中，确定每个集合是否是紧致的，是否为凸集合.

5. 考虑习题 3 中的集合.

6. 考虑习题 4 中的集合.

在习题 7~10 中，设 H 是过下列点的超平面.（a）求与超平面正交的向量 $\boldsymbol{n}$.（b）求线性函数 f 及实数 d 满足 $H=[f:d]$.

7. $\begin{bmatrix}1\\1\\3\end{bmatrix},\begin{bmatrix}2\\4\\1\end{bmatrix},\begin{bmatrix}-1\\-2\\5\end{bmatrix}$

8. $\begin{bmatrix}1\\-2\\1\end{bmatrix},\begin{bmatrix}4\\-2\\3\end{bmatrix},\begin{bmatrix}7\\-4\\4\end{bmatrix}$

9. $\begin{bmatrix}1\\0\\1\\0\end{bmatrix},\begin{bmatrix}2\\3\\1\\0\end{bmatrix},\begin{bmatrix}1\\2\\2\\0\end{bmatrix},\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$

10. $\begin{bmatrix}1\\2\\0\\0\end{bmatrix},\begin{bmatrix}2\\2\\-1\\-3\end{bmatrix},\begin{bmatrix}1\\3\\2\\7\end{bmatrix},\begin{bmatrix}3\\2\\-1\\-1\end{bmatrix}$

11. 设 $\boldsymbol{p}=\begin{bmatrix}1\\-3\\1\\2\end{bmatrix},\boldsymbol{n}=\begin{bmatrix}2\\1\\5\\-1\end{bmatrix},\boldsymbol{v}_1=\begin{bmatrix}0\\1\\1\\1\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}-2\\0\\1\\3\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}1\\4\\0\\4\end{bmatrix}$，且 H 是 $\mathbb{R}^4$ 中过点 $\boldsymbol{p}$ 且与 $\boldsymbol{n}$ 正交的超平面. $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 中哪些点与原点在 H 的同一侧，哪些点不如此.

12. $\boldsymbol{a}_1=\begin{bmatrix}2\\-1\\5\end{bmatrix},\boldsymbol{a}_2=\begin{bmatrix}3\\1\\3\end{bmatrix},\boldsymbol{a}_3=\begin{bmatrix}-1\\6\\0\end{bmatrix},\boldsymbol{b}_1=\begin{bmatrix}0\\5\\-1\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}1\\-3\\-2\end{bmatrix},$ $\boldsymbol{b}_3=\begin{bmatrix}2\\2\\1\end{bmatrix},\boldsymbol{n}=\begin{bmatrix}3\\1\\-2\end{bmatrix}$，且设 $A=\{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3\}$，$B=\{\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3\}$. 求与 $\boldsymbol{n}$ 正交，分割 A 与 B 的超平面 H. 存在平行于 H 的超平面严格分割 A 与 B 吗？

13. 设 $\boldsymbol{p}_1=\begin{bmatrix}2\\-3\\1\\2\end{bmatrix}$，$\boldsymbol{p}_2=\begin{bmatrix}1\\2\\-1\\3\end{bmatrix}$，$\boldsymbol{n}_1=\begin{bmatrix}1\\2\\4\\2\end{bmatrix}$，$\boldsymbol{n}_2=\begin{bmatrix}3\\3\\1\\5\end{bmatrix}$，且设 H_1 是 $\mathbb{R}^4$ 中的超平面，过点 $\boldsymbol{p}_1$ 且与 $\boldsymbol{n}_2$ 正交，H_2 是过点 $\boldsymbol{p}_2$ 且与 $\boldsymbol{n}_2$ 正交的超平面. 求 $H_1\cap H_2$ 的显式表达式（提示：求 $H_1\cap H_2$ 中一点 $\boldsymbol{p}$ 及两个线性无关向量 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 生成的平行于二维平面 $H_1\cap H_2$ 的子空间.）

14. 设 F_1,F_2 是 $\mathbb{R}^6$ 中的 4 维平面，假定 $F_1\cap F_2\neq\varnothing$，$F_1\cap F_2$ 可能是多少维的？

在习题 15~20 中，求线性函数 f 及特定的实数 d，使得$[f{:}d]$是习题中所描述的超平面 H.

15. 设 A 是 1×4 矩阵$[1\ \ -3\ \ 4\ \ -2]$且 b=5. 设 $H=\{\boldsymbol{x}\in\mathbb{R}^4: A\boldsymbol{x}=b\}$.

16. 设 A 是 1×5 矩阵$[2\ \ 5\ \ -3\ \ 0\ \ 6]$，注意 $\operatorname{Nul} A\in\mathbb{R}^5$. 令 H=Nul A.

17. 设 H 是 $\mathbb{R}^3$ 中的平面，由矩阵 $B=\begin{bmatrix}1&3&5\\0&2&4\end{bmatrix}$ 的行向量生成. 也就是说，H=Row B.（提示：H 与 Nul B 有何关系？见 6.1 节.）

18. 设 H 是 $\mathbb{R}^3$ 中的平面，由矩阵 $B=\begin{bmatrix}1&4&-5\\0&-2&8\end{bmatrix}$ 的行向量生成，即 H=Row B.

19. 设 H 是由矩阵 $B=\begin{bmatrix}1&0\\4&2\\-7&-6\end{bmatrix}$ 的列向量生成，即 H=Col B.（提示：Col B 与 Nul B^{T} 有何关系？见 6.1 节.）

20. 设 H 是由矩阵 $B=\begin{bmatrix}1&0\\5&2\\-4&-4\end{bmatrix}$ 的列向量生成，即 H=Col B.

在习题 21 和 22 中判断每个命题的真假，并说明理由.

21. a. 从 $\mathbb{R}$ 到 $\mathbb{R}^n$ 的线性变换称为线性函数.

    b. 若 f 是定义在 $\mathbb{R}^n$ 上的线性函数，则存在实数 k，使得对 $\mathbb{R}^n$ 中所有的 $\boldsymbol{x}$ 有 $f(\boldsymbol{x})=k\boldsymbol{x}$.

    c. 若超平面严格分割 A 与 B，则 $A\cap B=\varnothing$.

    d. 若 A 与 B 是闭凸集，且 $A\cap B=\varnothing$，则存在超平面严格分割 A 与 B.

22. a. 设 d 是实数且 f 是 $\mathbb{R}^n$ 上的非零线性函数，则 $[f:d]$ 是 $\mathbb{R}^n$ 上的超平面.
    b. 给定任意向量 $\boldsymbol{n}$ 及任意实数 d，则集合 $[\boldsymbol{x}:\boldsymbol{n}\cdot\boldsymbol{x}=d]$ 是超平面.
    c. 设 A 与 B 是非空的不相交集，且 A 是紧致的，B 是闭的，则存在超平面严格分割 A 与 B.
    d. 若存在超平面 H 不严格分割 A 与 B，则 $(\text{conv }A)\cap(\text{conv }B)\neq\varnothing$.
23. 设 $\boldsymbol{v}_1=\begin{bmatrix}1\\1\end{bmatrix}$, $\boldsymbol{v}_2=\begin{bmatrix}3\\0\end{bmatrix}$, $\boldsymbol{v}_3=\begin{bmatrix}5\\3\end{bmatrix}$, $\boldsymbol{p}=\begin{bmatrix}4\\1\end{bmatrix}$，求超平面 $[f:d]$ 严格分割 $\boldsymbol{p}$ 与凸集 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$.
24. 重复习题 23，这里 $\boldsymbol{v}_1=\begin{bmatrix}1\\2\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}5\\1\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}4\\4\end{bmatrix}$, $\boldsymbol{p}=\begin{bmatrix}2\\3\end{bmatrix}$.
25. 设 $\boldsymbol{p}=\begin{bmatrix}4\\1\end{bmatrix}$. 求超平面 $[f:d]$ 严格分割 $B(\boldsymbol{0},3)$ 与 $B(\boldsymbol{p},1)$.（提示：在求得 f 后，证明 $\boldsymbol{v}=(1-0.75)\boldsymbol{0}+0.75\boldsymbol{p}$ 既非在 $B(\boldsymbol{0},3)$ 也非在 $B(\boldsymbol{p},1)$ 中.）
26. 设 $\boldsymbol{q}=\begin{bmatrix}2\\3\end{bmatrix}$, $\boldsymbol{p}=\begin{bmatrix}6\\1\end{bmatrix}$. 求超平面 $[f:d]$ 严格分割 $B(\boldsymbol{q},3)$ 与 $B(\boldsymbol{p},1)$.
27. 给出 $\mathbb{R}^2$ 中一个闭子集 S 使得 conv S 不是闭的.
28. 给出 $\mathbb{R}^2$ 中一个例子，满足 A 是紧致的，B 是闭的且 $(\text{conv }A)\cap(\text{conv }B)\neq\varnothing$，但 A 与 B 不能被超平面严格分割.
29. 证明：开球 $B(\boldsymbol{p},\delta)=\{\boldsymbol{x}:\|\boldsymbol{x}-\boldsymbol{p}\|<\delta\}$ 是凸集.（提示：使用三角不等式.）
30. 证明：有界集的凸包是有界的.

练习题答案

首先，计算 $\boldsymbol{n}_1\cdot\boldsymbol{p}_1=-3,\boldsymbol{n}_2\cdot\boldsymbol{p}_2=7$，超平面 H_1 是方程 $x_1+x_2-2x_3=-3$ 的解集合，H_2 是方程 $-2x_1+x_2+3x_3=7$ 的解集合. 那么

$$H_1\cap H_2=\{\boldsymbol{x}:x_1+x_2-2x_3=-3,-2x_1+x_2+3x_3=7\}$$

这是 $H_1\cap H_2$ 的隐式表达式. 为了求显式表达式，通过行变换解方程组：

$$\begin{bmatrix}1&1&-2&-3\\-2&1&3&7\end{bmatrix}\sim\begin{bmatrix}1&0&-\frac{5}{3}&-\frac{10}{3}\\0&1&-\frac{1}{3}&\frac{1}{3}\end{bmatrix}$$

因此 $x_1=-\frac{10}{3}+\frac{5}{3}x_3$，$x_2=\frac{1}{3}+\frac{1}{3}x_3$, $x_3=x_3$. 令 $\boldsymbol{p}=\begin{bmatrix}-\frac{10}{3}\\\frac{1}{3}\\0\end{bmatrix},\boldsymbol{v}=\begin{bmatrix}\frac{5}{3}\\\frac{1}{3}\\1\end{bmatrix}$，通解可写成 $\boldsymbol{x}=\boldsymbol{p}+x_3\boldsymbol{v}$. 因此，$H_1\cap H_2$ 是过点 $\boldsymbol{p}$ 以 $\boldsymbol{v}$ 为方向的直线. 注意 $\boldsymbol{v}$ 与 $\boldsymbol{n}_1$ 和 $\boldsymbol{n}_2$ 正交.

8.5 多面体

本节学习一种重要的紧致凸集（即多面体）的几何性质. 这些集合产生于各类应用中，包括博弈论（9.1 节）、线性规划（9.2~9.4 节）及更一般的优化问题，如工程系统中反馈控制的设计.

$\mathbb{R}^n$ 中的**多面体**是一个有限点集的凸包. $\mathbb{R}^2$ 中的多面体就是平面上的多边形. 而 $\mathbb{R}^3$ 中的多面体就是我们所知的多面体. $\mathbb{R}^3$ 中的多面体的重要特征是它的面、棱及顶点. 例如，立方体有 6 个面、12 条棱、8 个顶点. 下面的定义为高维空间提供一些术语，当然也适用于 $\mathbb{R}^2$ 和 $\mathbb{R}^3$ 空间.

回顾 $\mathbb{R}^n$ 中集合的维数，它是包含该集合的维数最小的平面的维数. 注意：多面体是特殊的紧致凸集，因为 $\mathbb{R}^n$ 中的有限集是紧致的且它的凸集也是紧致的，由 8.4 节中的拓扑知识可知.

定义 设 S 是 $\mathbb{R}^n$ 中的紧致凸子集. S 的非空子集 F 称为 S 的（真）**面**，如果 $F \neq S$ 且存在超平面 $H = [f:d]$ 使得 $F = S \cap H$，以及要么 $f(s) \geqslant d$ 要么 $f(s) \leqslant d$. 超平面 H 称为 S 的**支撑超平面**. 若 F 的维数为 k，则 F 称为 S 的 k 面.

若 P 是 k 维的多面体，则 P 称为 **k 面体**. P 的 0 面称为**顶点**，1 面为**棱**，而（k−1）维面称为 S 的**面.**

例 1 设 S 是 $\mathbb{R}^3$ 中的立方体，当平面 H 只与立方体的表面接触而不穿过其内部，基于 H 的方向，$H \cap S$ 存在三种可能（见图 8-21）：

$H \cap S$ 可能是立方体的 2 维正方形面.

$H \cap S$ 可能是立方体的 1 维棱.

$H \cap S$ 可能是立方体的 0 维顶点.

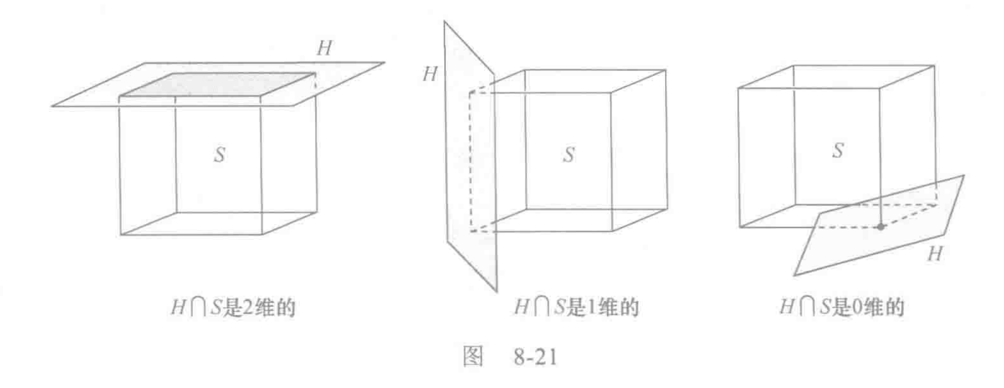


图 8-21

多面体的大部分应用与顶点有关. 因为它们具有如下定义中所描述的特殊性质.

定义 设 S 是凸集，若点 $\boldsymbol{p}$ 不在 S 中任何线段的内部，则称点 $\boldsymbol{p}$ 为 S 的**极端点**. 更为准确地，若 $\boldsymbol{x}, \boldsymbol{y} \in S$，且 $\boldsymbol{p} \in \overline{\boldsymbol{x}\boldsymbol{y}}$，则 $\boldsymbol{p} = \boldsymbol{x}$ 或 $\boldsymbol{p} = \boldsymbol{y}$. S 中所有极端点的集合称为 S 的**轮廓**.

任何紧凸集 S 的顶点都必然是 S 的极端点. 此结论在如下定理 14 的证明中证实. 如果所考虑的多面体可表示为 $\boldsymbol{p} = \operatorname{conv}\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\}, \boldsymbol{v}_1, \cdots, \boldsymbol{v}_k \in \mathbb{R}^n$，一般知道 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k$ 是 P 的极端点是有益的，然而它们可能包含多余的点. 例如，某些向量 $\boldsymbol{v}_i$ 可能是多面体的棱的中点. 当然，此时 $\boldsymbol{v}_i$ 在构造凸包时并不真正需要用到. 下面的定义描述了顶点的性质，使得它们成为极端点.

定义 集合 $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\}$ 是多面体 P 的**最小代表元**，若 $P = \operatorname{conv}\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\}$，且对每个 $i = 1, \cdots, k$，$\boldsymbol{v}_i \notin \operatorname{conv}\{\boldsymbol{v}_j : j \neq i\}$.

每个多面体都有一个最小代表元. 设 $P = \operatorname{conv}\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\}$，若某个 $\boldsymbol{v}_i$ 是其他点的凸组合，则 $\boldsymbol{v}_i$ 可从点集中消去而不改变凸包，重复这个过程，直到剩下最小代表元. 可以证明最小代表元是唯一的.

定理 14 $M=\{v_1,\cdots,v_k\}$ 是多面体 P 的最小代表元，则下面三个命题是等价的：

a. $\boldsymbol{p}\in M$.

b. $\boldsymbol{p}$ 是 P 的顶点.

c. $\boldsymbol{p}$ 是 P 的极端点.

证明 (a)$\Rightarrow$(b)，设 $\boldsymbol{p}\in M$ 且令 $Q=\text{conv}\ \{\boldsymbol{v}:\boldsymbol{v}\in M,\boldsymbol{v}\notin \boldsymbol{p}\}$. 由 M 的定义知，$\boldsymbol{p}\notin \text{conv}\,Q$，又因为 Q 是紧致的，由定理 13 知，存在超平面 H'严格分割$\{\boldsymbol{p}\}$与 Q. 设 H 是过点 $\boldsymbol{p}$ 且平行于 H'的超平面，见图 8-22.

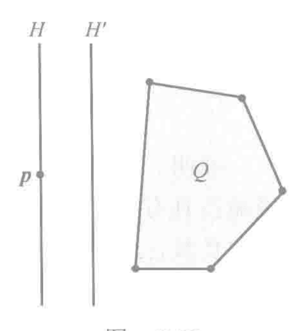


图 8-22

那么 Q 位于由 H 界定的闭的半空间 H^+的一侧，从而 $P\subseteq H^+$. 因而，H 在点 $\boldsymbol{p}$ 支撑 P. 进而 $\boldsymbol{p}$ 是 P 中位于 H 上的唯一点，故 $H\cap P=\{\boldsymbol{p}\}$，$\boldsymbol{p}$ 是 P 的顶点.

(b)$\Rightarrow$(c)，设 $\boldsymbol{p}$ 是 P 的顶点，则存在超平面 $H=[f{:}d]$，使得 $H\cap P=\{\boldsymbol{p}\}$ 且 $f(P)\geqslant d$，若 $\boldsymbol{p}$ 不是极端点，则存在 P 上的点 $\boldsymbol{x}$ 和 $\boldsymbol{y}$，使得 $\boldsymbol{p}=(1-c)\boldsymbol{x}+c\boldsymbol{y}$，$0<c<1$，即

$$c\boldsymbol{y}=\boldsymbol{p}-(1-c)\boldsymbol{x},\ \ \boldsymbol{y}=(\frac{1}{c})\boldsymbol{p}-(\frac{1}{c}-1)\boldsymbol{x}$$

于是满足 $f(\boldsymbol{y})=\left(\frac{1}{c}\right)f(\boldsymbol{p})-\left(\frac{1}{c}-1\right)f(\boldsymbol{x})$. 但 $f(\boldsymbol{p})=d$ 且 $f(\boldsymbol{x})\geqslant d$，故 $f(\boldsymbol{y})\leqslant\left(\frac{1}{c}\right)(d)-\left(\frac{1}{c}-1\right)\times(d)=d$.
另一方面，$\boldsymbol{y}\in P$，故 $f(\boldsymbol{y})\geqslant d$. 于是 $f(\boldsymbol{y})=d$，从而 $\boldsymbol{y}\in H\cap P$. 这与 $\boldsymbol{p}$ 是顶点矛盾，故 $\boldsymbol{p}$ 一定是极端点.（注意，这部分证明不需要 P 是多面体，它对任何紧凸集成立.）

(c)$\Rightarrow$(a)，显然 P 的任何极端点属于 M. ■

例 2 回顾集合 S 的轮廓是 S 的极端点的集合. 定理 14 证明了 $\mathbb{R}^2$ 中多边形的轮廓是顶点集.（见图 8-23）. 闭球的轮廓是它的边界. 开集没有极端点，故它的轮廓是空集. 闭的半空间没有极端点，故它的轮廓是空集. ■

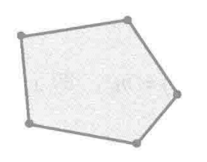
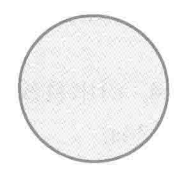
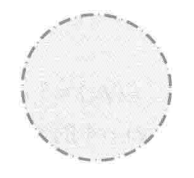
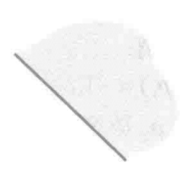
   

图 8-23

习题 18 要求证明凸集 S 中的点 $\boldsymbol{p}$ 是 S 的极端点当且仅当 $\boldsymbol{p}$ 从 S 中除去，剩下的点仍然构成一个凸集. 由此导出 S^*为 S 的任何子集且使得 conv S^*与 S 相等，则 S^*一定包含 S 的轮廓. 例 2 中的集合表明了通常 S^*可能要比 S 的轮廓大些. 然而，正如定理 15 要证明的，当 S 是紧致凸集时，我们可以取 S^* 只是 S 的轮廓就够了. 因而，在满足其凸包等于 S 的条件下，每个非空紧（凸）集 S 有极端点，所有极端点的集合就是 S 的最小子集.

定理 15 设 S 为非空紧凸集，则 S 是它的轮廓（S 的极端点的集合）的凸包.

证明 对 S 的维数运用归纳法证明[㊀]. ■

定理 15 的一个重要应用是下面的定理，它是线性规划发展中关键的理论成果之一. 线性函数是连续的，而连续函数在紧集上总可以取得最大值和最小值. 定理 16 的意义在于，对紧凸集来说，最大值（最小值）实际上可以在 S 的极端点取得.

定理 16 设 f 是定义在非空的紧凸集 S 上的线性函数. 则存在 S 的极端点 $\hat{\boldsymbol{v}}$ 和 $\hat{\boldsymbol{w}}$，使得

$$f(\hat{\boldsymbol{v}})=\max_{\boldsymbol{v}\in S} f(\boldsymbol{v}) \text{ 且 } f(\hat{\boldsymbol{w}})=\min_{\boldsymbol{v}\in S} f(\boldsymbol{v})$$

证明 假定 f 在 S 的某点 $\boldsymbol{v}'$ 达到 S 上的最大值 m，即 $f(\boldsymbol{v}')=m$. 我们希望证明 S 存在一个极端点具有相同的性质. 由定理 15 知，$\boldsymbol{v}'$ 是 S 的极端点和凸组合，即存在 S 中的极端点 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k$ 及非负数 $c_1,\cdots,c_k$，使得

$$\boldsymbol{v}'=c_1\boldsymbol{v}_1+\cdots+c_k\boldsymbol{v}_k \text{ 且 } c_1+\cdots+c_k=1$$

若 S 中没有极端点满足 $f(\boldsymbol{v})=m$，则

$$f(\boldsymbol{v}_i)<m\text{，这里 } i=1,\cdots,k$$

因为 m 是 f 在 S 上的最大值. 那么由于 f 是线性的，

$$\begin{aligned} m=f(\boldsymbol{v}')&=f(c_1\boldsymbol{v}_1+\cdots+c_k\boldsymbol{v}_k)=c_1f(\boldsymbol{v}_1)+\cdots+c_kf(\boldsymbol{v}_k)\\ &<c_1m+\cdots+c_km=m(c_1+\cdots+c_k)=m \end{aligned}$$

这个矛盾表明 S 的某个极端点 $\hat{\boldsymbol{v}}$ 满足 $f(\hat{\boldsymbol{v}})=m$.

对于 $\hat{\boldsymbol{w}}$ 的证明类似. ■

例 3 设 $\mathbb{R}^2$ 中的点 $\boldsymbol{p}_1=\begin{bmatrix}-1\\0\end{bmatrix},\boldsymbol{p}_2=\begin{bmatrix}3\\1\end{bmatrix},\boldsymbol{p}_3=\begin{bmatrix}1\\2\end{bmatrix}$，且 $S=\operatorname{conv}\{\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3\}$. 对下面的每个线性函数 f，求 f 在 S 上的最大值 m，求 S 上所有满足 $f(\boldsymbol{x})=m$ 的点 $\boldsymbol{x}$.

a. $f_1(x_1,x_2)=x_1+x_2$　　b. $f_2(x_1,x_2)=-3x_1+x_2$　　c. $f_3(x_1,x_2)=x_1+2x_2$

解 由定理 16，最大值在 S 的某个极端点取得. 故为了求得 m，算出 f 在每个极端点的值，并选出最大值.

a. $f_1(\boldsymbol{p}_1)=-1$，$f_1(\boldsymbol{p}_2)=4$，$f_1(\boldsymbol{p}_3)=3$，故 $m_1=4$. 画出直线 $f_1(x_1,x_2)=m_1$，即 $x_1+x_2=4$，并注意到 $\boldsymbol{x}=\boldsymbol{p}_2$ 是 S 上唯一使得 $f_1(\boldsymbol{x})=4$ 的点. 见图 8-24a.

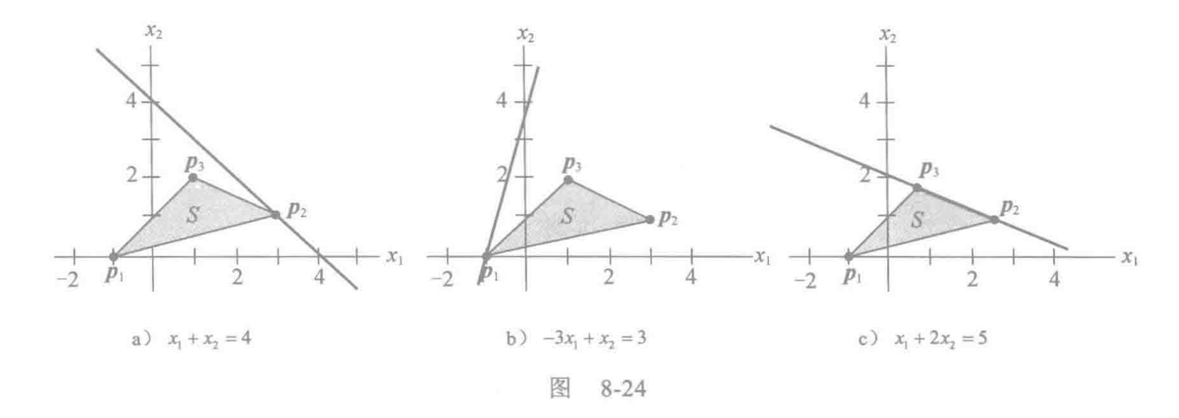
b. $f_2(\boldsymbol{p}_1)=3$，$f_2(\boldsymbol{p}_2)=-8$，$f_2(\boldsymbol{p}_3)=-1$，故 $m_2=3$. 画出直线 $f_2(x_1,x_2)=m_2$，即 $-3x_1+x_2=3$，并注意到 $\boldsymbol{x}=\boldsymbol{p}_1$ 是 S 上唯一使得 $f_2(\boldsymbol{x})=3$ 的点. 见图 8-24b.

c. $f_3(\boldsymbol{p}_1)=-1$，$f_3(\boldsymbol{p}_2)=5$，$f_3(\boldsymbol{p}_3)=5$，故 $m_3=5$. 画出直线 $f_3(x_1,x_2)=m_3$，即 $x_1+2x_2=5$. 这里 f_3 在 $\boldsymbol{p}_2,\boldsymbol{p}_3$ 取得最大值，从而在 $\boldsymbol{p}_2$ 和 $\boldsymbol{p}_3$ 的凸包上的每一点取得最大值. 见图 8-24c. ■

例 3 中对于 $\mathbb{R}^2$ 空间说明的情形也适用于高维空间. 线性函数 f 在多面体 P 上的最大值在支

㊀ 详细证明可以参考 Steven R. Lay 的 *Convex Sets and Their Applications*(New York: John Wiley & Sons, 1982; Mineola, NY:Dover Publications, 2007), p.43.

撑超平面与 P 的交点处取得. 这个交点要么是 P 的单个极端点，要么是 P 的 2 个或更多极端点的凸包. 在每种情况下，交集都是一个多面体，且它的极端点构成了 P 的极端点的子集.



a) $x_1+x_2=4$　　b) $-3x_1+x_2=3$　　c) $x_1+2x_2=5$

图 8-24

由定义，多面体是有限点集的凸包. 这是多面体的显式表达式，因为它指定了集合中的所有点. 多面体也可以用隐式表示为有限个闭的半空间的交集. 例 4 说明了 $\mathbb{R}^2$ 空间中的这个情形.

例 4　设 $\mathbb{R}^2$ 中的点 $\boldsymbol{p}_1=\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{p}_2=\begin{bmatrix}1\\0\end{bmatrix}, \boldsymbol{p}_3=\begin{bmatrix}3\\2\end{bmatrix}$, 且 $S=\text{conv}\{\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3\}$. 由简单的代数可知，过点 $\boldsymbol{p}_1$ 和 $\boldsymbol{p}_2$ 的直线为 $x_1+x_2=1$，且 S 在满足下列条件的直线的一侧:

$$x_1+x_2\geqslant 1，\text{或等价地，}\ -x_1-x_2\leqslant -1$$

类似地，过点 $\boldsymbol{p}_2$ 和 $\boldsymbol{p}_3$ 的直线为 $x_1-x_2=1$，且 S 在满足下面条件的直线一侧:

$$x_1-x_2\leqslant 1$$

同样，过点 $\boldsymbol{p}_3$ 和 $\boldsymbol{p}_1$ 的直线为 $-x_1+3x_2=3$，且 S 在满足下面条件的直线一侧:

$$-x_1+3x_2\leqslant 3$$

见图 8-25. 于是 S 可视为下面线性不等式组的解集:

$$\begin{aligned}-x_1-x_2&\leqslant -1\\x_1-x_2&\leqslant 1\\-x_1+3x_2&\leqslant 3\end{aligned}$$

这个线性不等式组可以写成 $A\boldsymbol{x}\leqslant\boldsymbol{b}$，其中

$$A=\begin{bmatrix}-1&-1\\1&-1\\-1&3\end{bmatrix},\ \boldsymbol{x}=\begin{bmatrix}x_1\\x_2\end{bmatrix},\ \text{且}\boldsymbol{b}=\begin{bmatrix}-1\\1\\3\end{bmatrix}$$

注意两个向量之间的不等号，例如 $A\boldsymbol{x}$ 与 $\boldsymbol{b}$ 之间的不等号是用于两个向量每一组对应分量之间的. ■

在简单的情形中，图像法求解是可行的. 下面的例子说明了当图像中的几个我们感兴趣的点靠得很近而无法确认时的处理方法.

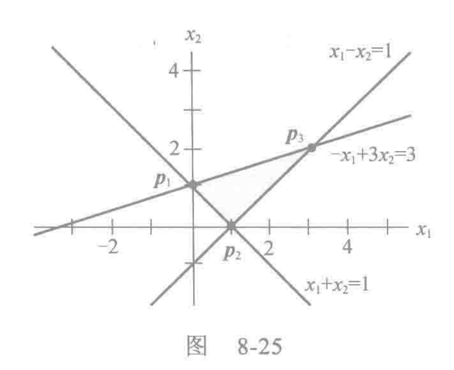


图 8-25

例 5 设 P 是 $\mathbb{R}^2$ 中的点集且满足 $A\boldsymbol{x}\leqslant\boldsymbol{b}$，这里

$$A=\begin{bmatrix}1 & 3\\1 & 1\\3 & 2\end{bmatrix}，且\boldsymbol{b}=\begin{bmatrix}18\\8\\21\end{bmatrix}$$

且 $\boldsymbol{x}\geqslant\boldsymbol{0}$．求 P 的最小代表元．

解 条件 $\boldsymbol{x}\geqslant\boldsymbol{0}$ 说明 P 在 $\mathbb{R}^2$ 中的第一象限，是线性规划问题中的典型条件，$A\boldsymbol{x}\leqslant\boldsymbol{b}$ 中的三个不等式涉及三条边界线：

$$（1）\ x_1+3x_2=18\ （2）\ x_1+x_2=8\ （3）\ 3x_1+2x_2=21$$

三条直线的斜率都是负的，故 P 的大体形状是很容易想象的．大体勾画出的三条直线将表明(0,0),(7,0)(0,6)为多面体 P 的顶点．

直线（1）、（2）和（3）的交点情况如何？有时候从图形中明显地知道有哪些交点．但如果不明显时，如下的代数过程会有作用：

当找到对应于两个不等式的交点时，在其他不等式中检验看此点是否在多面体中．

（1）和（2）的交点是 $\boldsymbol{p}_{12}=(3,5)$，两个坐标点都是非负的，故 $\boldsymbol{p}_{12}$ 满足除第三个不等式的所有不等式．检验得，

$$3(3)+2(5)=19<21$$

交点满足不等式（3），故 $\boldsymbol{p}_{12}$ 在多面体中．

（2）和（3）的交点是 $\boldsymbol{p}_{23}=(5,3)$．故 $\boldsymbol{p}_{23}$ 满足除第一个不等式的所有不等式．检验得，

$$1(5)+3(3)=14<18$$

说明 $\boldsymbol{p}_{23}$ 在多面体中．

最后（1）和（3）的交点是 $\boldsymbol{p}_{13}=\left(\dfrac{27}{7},\dfrac{33}{7}\right)$，在不等式（2）中．检验得，

$$1\left(\frac{27}{7}\right)+1\left(\frac{33}{7}\right)=\frac{60}{7}\approx 8.6>8$$

$\boldsymbol{p}_{13}$ 不满足第二个不等式，表明 $\boldsymbol{p}_{13}$ 不在 P 中．结论是多面体 P 的最小代表元是

$$\left\{\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}7\\0\end{bmatrix},\begin{bmatrix}3\\5\end{bmatrix},\begin{bmatrix}5\\3\end{bmatrix},\begin{bmatrix}0\\6\end{bmatrix}\right\}$$

■

本节其余部分将讨论 $\mathbb{R}^3$（和高维空间）中两类基本多面体的构造. 第一类出现在线性规划问题中. 两类多面体为“可视化” $\mathbb{R}^4$ 空间提供了一种著名的方法.

单纯形

单纯形是仿射不相关的有限向量集的凸包. 为了构造一个 k 维的单纯形（或 k-单纯形），过程如下：

0 维单纯形 S^0：单个点$\{\boldsymbol{v}_1\}$

1 维单纯形 S^1：$\text{conv}\left(S^0 \cup \{\boldsymbol{v}_2\}\right)$，$\boldsymbol{v}_2$ 不在仿射集 S^0 中

2 维单纯形 S^2：$\text{conv}\left(S^1 \cup \{\boldsymbol{v}_3\}\right)$，$\boldsymbol{v}_3$ 不在仿射集 S^1 中

$\vdots$

k 维单纯形 S^k：$\text{conv}\left(S^{k-1} \cup \{\boldsymbol{v}_{k+1}\}\right)$，$\boldsymbol{v}_{k+1}$ 不在仿射集 S^{k-1} 中

单纯形 S^1 是一线段. 三角形 S^2 由 S^1 及不在包含此线段的直线上的一点 $\boldsymbol{v}_3$ 一起构作凸包而得.（见图 8-26.）四面体 S^3 由 S^2 及不在包含 S^2 的平面上的一点 $\boldsymbol{v}_4$ 一起构作凸包而得.

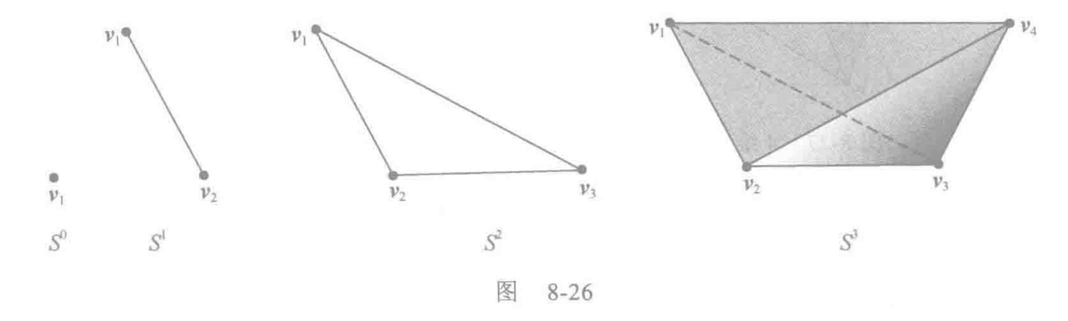


图 8-26

在继续讲述之前，我们考察一下正在呈现的一些模式. 三角形 S^2 有三条边，每条边是一条线段，如 S^1. 这三条线段是从哪里来呢？其中一条是 S^1，另外两条中的一条来自端点在 $\boldsymbol{v}_2$ 与新点 $\boldsymbol{v}_3$ 的连线. 第三条来自另一端点 $\boldsymbol{v}_1$ 与新点 $\boldsymbol{v}_3$ 的连线. 你也可以把这个形成过程说成是把 S^1 的每个端点延伸（或拉伸）到 S^1 所在的直线之外的新点 $\boldsymbol{v}_3$ 而得到这两条线段的.

图 8-26 中的四面体 S^3 有 4 个三角形面. 其中一个是原三角形面 S^2，而其他三个来自向外延伸 S^2 的三条边至新点 $\boldsymbol{v}_4$. 也要注意 S^2 的三个顶点同时被拉伸形成 S^3 的三条棱，而 S^3 的其他三条棱来自 S^2 的边. 这个观点启发我们如何去“可视化”4 维单纯形 S^4.

4 维单纯形 S^4 是由 S^3 与 S^3 所在的三维空间之外的一点 $\boldsymbol{v}_5$ 构作凸包而得的. 在三维空间中画出一个完整的图当然是不可能的，但图 8-27 是值得参考的：S^4 有 5 个顶点，其中任何 4 个顶点确定 S^4 的一个面，这些面都形如四面体 S^3. 例如，图中侧重显示了顶点分别为 $\boldsymbol{v}_1$，$\boldsymbol{v}_2$，$\boldsymbol{v}_4$，$\boldsymbol{v}_5$ 和 $\boldsymbol{v}_2$，$\boldsymbol{v}_3$，$\boldsymbol{v}_4$，$\boldsymbol{v}_5$ 的两个面（四面体）. 单纯形 S^4 共有 5 个这样的面（四面体）. 图 8-27 给定了 S^4 的 10 条棱，通过图中这些棱也“可视化”了 S^4 的 10 个 2-面三角形面.

图 8-28 展示了 4 维单纯形 S^4 的另一种表示，这次第 5 个顶点在四面体 S^3 的“里面”，高亮显示的面（四面体）也在四面体 S^3 的“里面”.

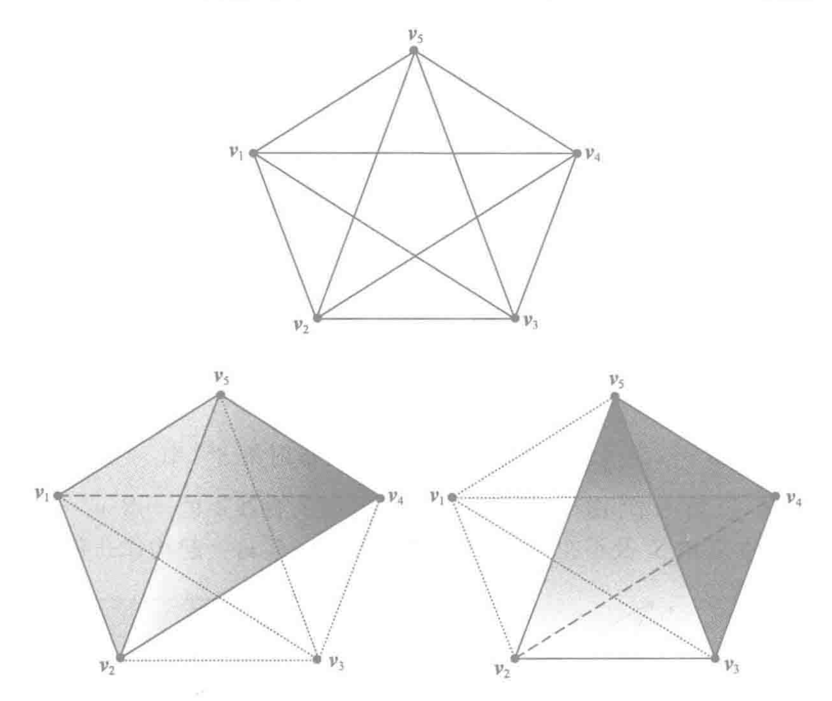


图 8-27　4 维单纯形投影到 $\mathbb{R}^2$ 上，侧重显示四面体的两个面

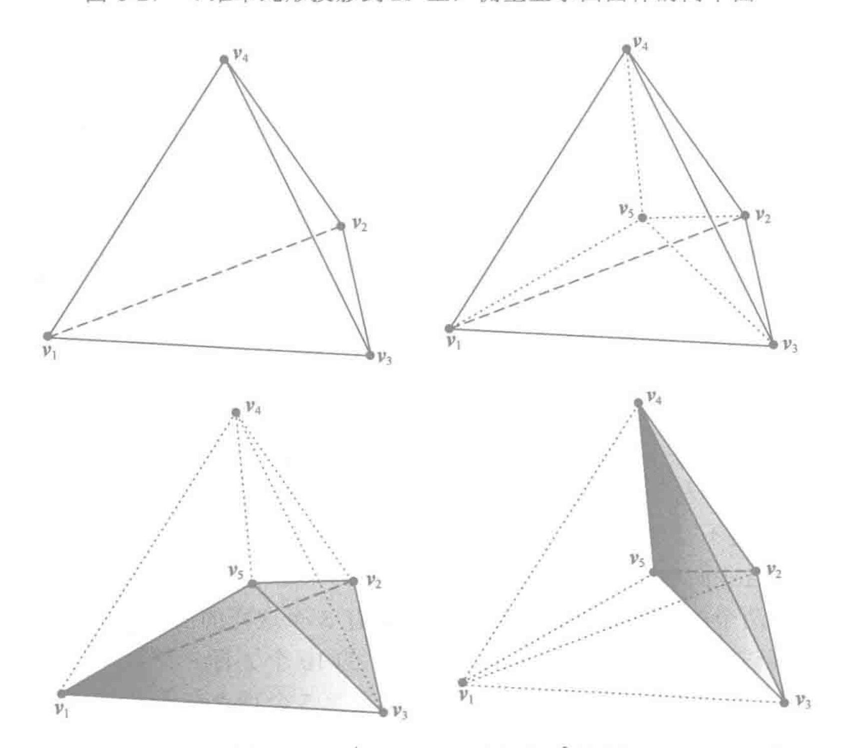


图 8-28　S^4 的第 5 个顶点在 S^3 里面

超立方体

设 $I_i=\overline{\mathbf{0}e_i}$ 是从原点 $\mathbf{0}$ 到 $\mathbb{R}^n$ 中标准基向量 e_i 的线段. 那么对于 $k(1\leqslant k\leqslant n)$，向量和[㊀]

$$C^k=I_1+I_2+\cdots I_k$$

称为 k 维**超立方体**.

为了看见 C^k 的构造，从简单的例子开始. 超立方体 C^1 就是线段 I_1. 若 C^1 沿 e_2 平移，由起始位置与最后位置一起构作凸包就描述了正方形 C^2（见图 8-29）. C^2 沿 e_3 平移就构造出立方体 C^3. 类似地，C^3 沿向量 e_4 平移就构造出超立方体 C^4.

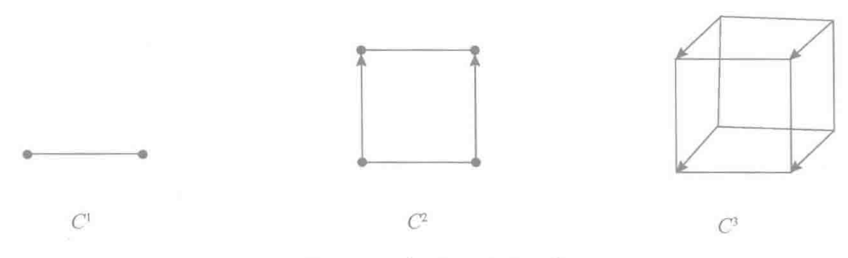


图 8-29 构造立方体 C^3

同样，对四维情况的 C^4，这是很难想象的，但图 8-30 显示了 C^4 的二维投影. 拉伸 C^3 的每一条棱进入 C^4 的一个正方形面中，这个过程可视为形成 C^4 的一个正方形面（2-面）的过程. 类似地，拉伸 C^3 的每个面进入 C^4 的立方体面（3-面）中，这个过程视为形成 C^4 的一个立方体面（3-面）的过程. 图 8-31 展示了 C^4 的三个面（3-面）. 图 8-31a 的高亮部分显示了来自 C^3 左边正方形面的立方体. 图 8-31b 显示了来自 C^3 前面正方形面的立方体. 图 8-31c 显示了来自 C^3 顶部正方形面的立方体.

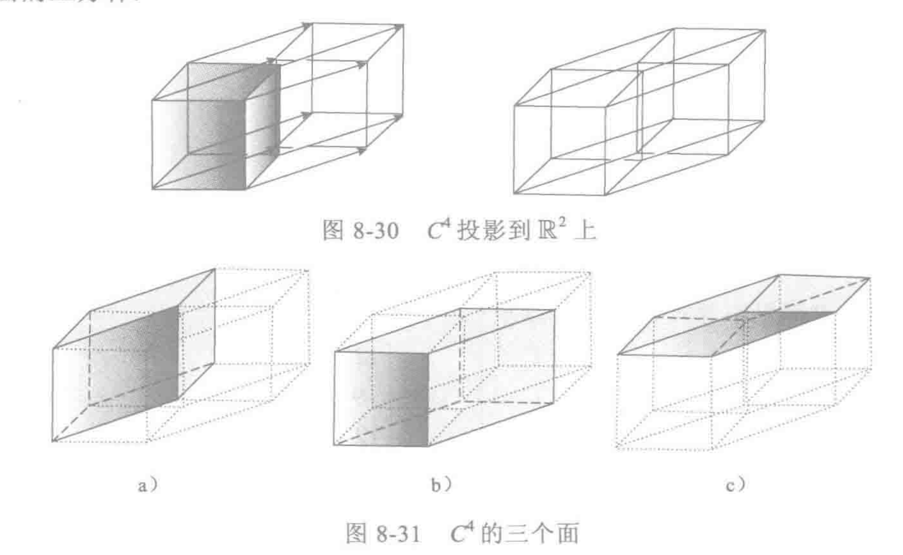


图 8-30 C^4 投影到 $\mathbb{R}^2$ 上

图 8-31 C^4 的三个面

㊀ 集合 A 和 B 的向量和定义为 $A+B=\{c:c=a+b,a\in A,b\in B\}$.

图 8-32 展示了 C^4 的另一种表示，在这里平移后的立方体位于 C^3 “里面”．这种方式比较容易“看到” C^4 的立方体面（3-面），因为没有干扰因素．

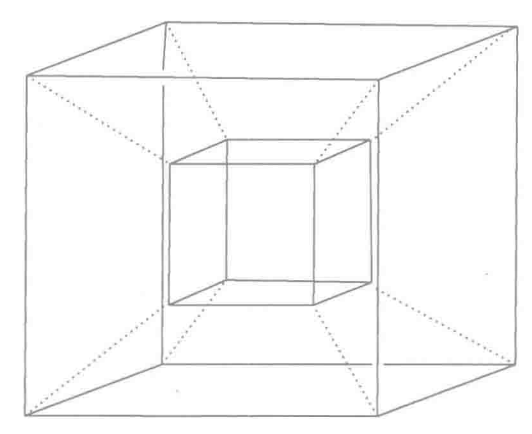


图 8-32 平移 C^3 的图像到 C^3 “里面”获得 C^4

总之，4 维超立方体 C^4 有 8 个立方体面（3-面）．其中两个就是原始的 C^3 及平移后的 C^3，其他 6 个来自 C^3 的 6 个正方形面拉伸（缩）进入 C^4 的立方体面（3-面）所形成. C^4 的 2-面来自原始的 C^3 的 6 个正方形面和平移后的 C^3 的 6 个正方形面，以及原始的 C^3 的 12 条棱拉伸（缩）进入 C^4 的正方形面（2-面）所形成. 因而有 2×6+12=24 个正方形面（2-面）．为了计算棱数，把 C^3 的棱数乘以 2 再加上 C^3 的顶点数. 于是 C^4 的棱数为 2×12+8=32. C^4 中的顶点来自 C^3 及其平移后的 C^3 的顶点，故有 2×8=16 个顶点.

多面体研究中真正著名的结论之一是下面的公式，其首次证明由欧拉（Leonard Euler，1707—1783）给出. 它为多面体中不同维数的面的个数建立了一个简单的关系. 为了简化公式，用 $f_k(P)$ 记 n 维多面体 P 中 k 维面的个数. ㊀ 那么有

$$\text{欧拉公式：}\quad \sum_{k=0}^{n-1}(-1)^k f_k(P) = 1+(-1)^{n-1}$$

特别地，当 $n=3$ 时，$v-e+f=2$，这里 v，e 和 f 分别为 P 的顶点、棱和面的个数.

练习题

1. 设多面体 P 满足不等式 $A\boldsymbol{x} \leqslant \boldsymbol{b}$，这里

$$A=\begin{bmatrix}1 & 3\\1 & 2\\2 & 1\end{bmatrix},\ \text{且}\boldsymbol{b}=\begin{bmatrix}12\\9\\12\end{bmatrix}$$

且 $\boldsymbol{x} \geqslant \boldsymbol{0}$. 求 P 的最小代表元.

㊀ 证明参见 Steven R. Lay 的 *Convex Sets and Their Applications* (New York: John Wiley & Sons, 1982; Mineola, NY: Dover Publications, 2007), p.131.

习题 8.5

1. 设 $\mathbb{R}^2$ 中的点 $\boldsymbol{p}_1=\begin{bmatrix}1\\0\end{bmatrix},\boldsymbol{p}_2=\begin{bmatrix}2\\3\end{bmatrix},\boldsymbol{p}_3=\begin{bmatrix}-1\\2\end{bmatrix}$ 且 $S=\text{conv}\{\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3\}$. 对下面的每个线性函数 f，求 f 在 S 上的最大值 m，求 S 上所有满足 $f(\boldsymbol{x})=m$ 的点 $\boldsymbol{x}$.
   a. $f(x_1,x_2)=x_1-x_2$　　b. $f(x_1,x_2)=x_1+x_2$
   c. $f(x_1,x_2)=-3x_1+x_2$
2. 设 $\mathbb{R}^2$ 中的点 $\boldsymbol{p}_1=\begin{bmatrix}0\\-1\end{bmatrix},\boldsymbol{p}_2=\begin{bmatrix}2\\1\end{bmatrix},\boldsymbol{p}_3=\begin{bmatrix}1\\2\end{bmatrix}$ 且 $S=\text{conv}\{\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3\}$. 对下面的每个线性函数 f，求 f 在 S 上的最大值 m，求 S 上所有满足 $f(\boldsymbol{x})=m$ 的点 $\boldsymbol{x}$.
   a. $f(x_1,x_2)=x_1+x_2$　　b. $f(x_1,x_2)=x_1-x_2$
   c. $f(x_1,x_2)=-2x_1+x_2$
3. 重复习题 1，这里取 m 是 f 在 S 上的最小值.
4. 重复习题 2，这里取 m 是 f 在 S 上的最小值.

在习题 5~8 中，求由不等式 $A\boldsymbol{x}\leqslant\boldsymbol{b}$ 且 $\boldsymbol{x}\geqslant 0$ 给定的多面体的最小代表元.

5. $A=\begin{bmatrix}1&2\\3&1\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}10\\15\end{bmatrix}$
6. $A=\begin{bmatrix}2&3\\4&1\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}18\\16\end{bmatrix}$
7. $A=\begin{bmatrix}1&3\\1&1\\4&1\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}18\\10\\28\end{bmatrix}$
8. $A=\begin{bmatrix}2&1\\1&1\\1&2\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}8\\6\\7\end{bmatrix}$
9. 设 $S=\{(x,y):x^2+(y-1)^2\leqslant 1\}\cup\{(3,0)\}$. 原点是凸集 S 的极端点吗？原点是凸集 S 的顶点吗？
10. 求 $\mathbb{R}^2$ 中闭凸集 S，使得它的轮廓 P 是非空集但 $\text{conv}\,P\neq S$.
11. 求 $\mathbb{R}^2$ 中有界凸集 S，使得它的轮廓 P 是非空集但凸集 $\text{conv}\,P\neq S$.
12. a. 确定 5 维单纯形 S^5 中 k 面的个数，这里 $k=0,1,\cdots,4$. 证明你的答案满足欧拉公式.
    b. 制作一张 $f_k(S^n)$ 值的图表，这里 $n=1,\cdots,5$ 且 $k=0,1,\cdots,4$. 你可以看出规律吗？猜测 $f_k(S^n)$ 的一个一般公式.
13. a. 确定 5 维超立方体 C^5 中 k 面的个数，这里 $k=0,1,\cdots,4$. 证明你的答案满足欧拉公式.
    b. 制作一张 $f_k(C^n)$ 值的图表，这里 $n=1,\cdots,5$ 且 $k=0,1,\cdots,4$. 你可以看出规律吗？猜测 $f_k(C^n)$ 的一个一般公式.
14. 设 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k$ 是 $\mathbb{R}^n$ 中线性无关的向量 $(1\leqslant k\leqslant n)$. 那么集合 $X^k=\text{conv}\{\pm\boldsymbol{v}_1,\cdots,\pm\boldsymbol{v}_k\}$ 称为 **k 维交叉多面体**.
    a. 画出 X^1，X^2.
    b. 确定 3 维交叉多面体 X^3 中 k 面的个数，这里 $k=0,1,2$. X^3 另外的名称是什么？
    c. 确定 4 维交叉多面体 X^4 中 k 面的个数，这里 $k=0,1,2,3$. 证明你的答案满足欧拉公式.
    d. 求 $f_k(X^n)$ 的公式—— X^n 中 k 面的个数，$0\leqslant k\leqslant n-1$.
15. **k-金字塔** P^k 是 $k-1$ 维多面体 Q 与点 $x\notin\text{aff}\,Q$ 的凸包. 求下列公式，用 $f_j(Q)(j=0,\cdots,n-1)$ 表达.
    a. P^n 的顶点个数：$f_0(P^n)$.
    b. P^n 中 k-面的个数：$f_k(P^n)$，$1\leqslant k\leqslant n-2$.
    c. P^n 中 $n-1$ 维面的个数：$f_{n-1}(P^n)$.

在习题 16 和 17 中判断正误，并证明你的答案.

16. a. 多面体是有限点集的凸包.
    b. 设 $\boldsymbol{p}$ 是凸集 S 的极端点. 若 $\boldsymbol{u},\boldsymbol{v}\in S$，$\boldsymbol{p}\in\overline{\boldsymbol{uv}}$，且 $\boldsymbol{p}\neq\boldsymbol{u}$，那么 $\boldsymbol{p}=\boldsymbol{v}$.
    c. 若 S 是 $\mathbb{R}^n$ 上的非空凸子集，则 S 是其轮廓的凸包.
    d. 4 维单纯形 S^4 有 5 个面，每个是一个 3 维的四面体.
17. a. $\mathbb{R}^3$ 中的立方体有 5 个面（2-面）.
    b. 点 $\boldsymbol{p}$ 是多面体 P 的极端点当且仅当它是 P 的顶点.

c. 若 S 是非空紧凸集，且有线性函数在点 $\boldsymbol{p}$ 达到最大值，则 $\boldsymbol{p}$ 是 S 的极端点.

d. 2 维多面体总是具有相同的顶点数及棱数.

18. 设 $\boldsymbol{v}$ 是凸集 S 中的一元素. 证明：$\boldsymbol{v}$ 是 S 的极端点当且仅当集合 $\{\boldsymbol{x}\in S:\boldsymbol{x}\neq\boldsymbol{v}\}$ 是凸集.

19. 若 $c\in\mathbb{R}$ 且 S 是一集合，定义 $cS=\{c\boldsymbol{x}:\boldsymbol{x}\in S\}$. 设 S 是凸集且 $c>0$，$d>0$. 证明：$cS+dS=(c+d)S$.

20. 举例以证明习题 19 中 S 的凸性是必要的.

21. 若 A 和 B 是凸集，证明 $A+B$ 是凸的

22. 多面体（3 维的）称为**正多面体**，若它所有的面是全等正多边形且在顶点处的所有角都是相等的. 利用以下细节证明仅有 5 种正多面体.

a. 设正多面体有 r 个面，每个面是 k 边正多边形，且 s 条边在每个顶点相交. 用 v 和 e 分别记多面体的顶点数和棱数，证明 $kr=2e$，$sv=2e$.

b. 使用欧拉公式证明 $\frac{1}{s}+\frac{1}{k}=\frac{1}{2}+\frac{1}{e}$.

c. 求（b）中方程满足问题的几何约束的所有整数解.（k 与 s 可以有多少？）

下面的信息供参考：5 个正多面体是四面体（4，6，4），立方体（8，12，6），八面体（6，12，8），十二面体（20，30，12），二十面体（12，30，20）.（括号里的数字分别表示顶点数、棱数、面数.）

练习题答案

1. 矩阵不等式 $A\boldsymbol{x}\leqslant\boldsymbol{b}$ 产生如下不等式组：

$$\text{(a)}\quad x_1+3x_2\leqslant 12$$
$$\text{(b)}\quad x_1+2x_2\leqslant 9$$
$$\text{(c)}\quad 2x_1+x_2\leqslant 12$$

条件 $\boldsymbol{x}\geqslant\boldsymbol{0}$ 说明多面体在平面的第一象限，一个顶点是（0，0）. 当 $x_2=0$ 时，三条直线在 x_1 上的截距是 12，9 和 6，故（6，0）是顶点. 当 $x_1=0$ 时，三条直线在 x_2 上的截距是 4，4.5 和 12，故（0，4）是顶点.

三条直线如何相交于 x_1 和 x_2 的正值？（a）与（b）的交点是 $\boldsymbol{p}_{ab}=(3,3)$. 在（c）上检验 $\boldsymbol{p}_{ab}$ 得 $2(3)+1(3)=9<12$，故 $\boldsymbol{p}_{ab}$ 在 P 中.（b）与（c）的交点是 $\boldsymbol{p}_{bc}=(5,2)$. 在（a）上检验 $\boldsymbol{p}_{bc}$ 得 $1(5)+3(2)=11<12$，故 $\boldsymbol{p}_{bc}$ 在 P 中.（a）与（c）的交点是 $\boldsymbol{p}_{ac}=(4.8,2.4)$. 在（b）上检验 $\boldsymbol{p}_{ac}$ 得 $1(4.8)+2(2.4)=9.6>9$，故 $\boldsymbol{p}_{ac}$ 不在 P 中.

最后，多面体（多边形）的五个顶点（极端点）是（0，0），（6，0），（5，2），（3，3）和（0，4），这些点构成 P 的最小代表元. 如图 8-33 所示.

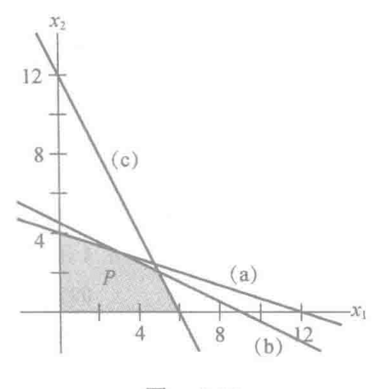


图 8-33

8.6 曲线与曲面

几千年以来，设计者们用长细木条建造了一艘艘船体. 近代，人们用长而柔韧的金属条架构出汽车和飞机的表面. 重物与钉子把条片塑造成称为*天然三次样条*的光滑曲线. 两个相邻控制点间（钉子或重物）的曲线有一个三次多项式的参数表达式. 遗憾的是，由于钉子与重物在条片上具有作用力，这种曲线在控制点稍有位移时会影响整条曲线的形状. 设计工程师们早就想局部地控制曲线——在这种控制中当移动一个控制点时仅影响曲线的一小部分. 1962 年，法国汽车工程师皮埃尔 • 贝塞尔通过增加额外的控制点与使用由他名字命名的一类曲线解决了这个问题.

贝塞尔曲线

如下所描述的曲线不仅在工程学而且在计算机图形学中都有着重要的应用. 例如，它们可应用于 Adobe Illustrator 与 Macromedia Freehand 中，也可应用于程序设计语言中，如 OpenGL. 这些曲线允许程序用相对少量的控制点来存储有关曲线段和表面的确切信息. 对曲线段和表面的所有图形命令只需要用来计算控制点. 曲线的特殊结构使之能通过“图形管道”加速其他计算，并把计算的终端显示在屏幕上.

8.3 节的习题介绍了二次贝塞尔曲线，并且给出了构造高次贝塞尔曲线的一种方法. 本节讨论二次和三次贝塞尔曲线，它们由三个或四个控制点 $\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3$ 确定. 这些点可以是 $\mathbb{R}^2$ 或 $\mathbb{R}^3$ 空间的点，亦可用齐次形式表示为 $\mathbb{R}^3$，$\mathbb{R}^4$ 空间中的点. 这些曲线的标准表达式为：

$$\boldsymbol{w}(t) = (1-t)^2 \boldsymbol{p}_0 + 2t(1-t)\boldsymbol{p}_1 + t^2 \boldsymbol{p}_2 \tag{1}$$

$$\boldsymbol{x}(t) = (1-t)^3 \boldsymbol{p}_0 + 3t(1-t)^2 \boldsymbol{p}_1 + 3t^2(1-t)\boldsymbol{p}_2 + t^3 \boldsymbol{p}_3 \tag{2}$$

其中 $0 \leqslant t \leqslant 1$.

图 8-34 给出两类典型的曲线. 通常，曲线仅过起始与最终的控制点，但贝塞尔曲线总是位于控制点的凸包里.（见 8.3 节的习题 21~24.）

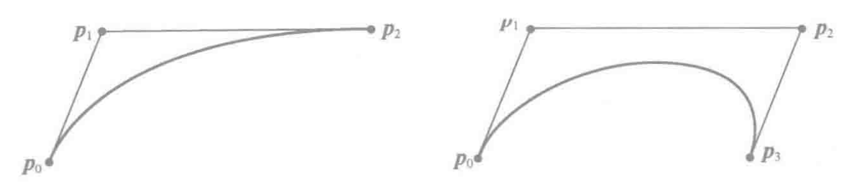


图 8-34　二次和三次贝塞尔曲线

贝塞尔曲线之所以在计算机图形学中非常有用，这是由于它们的基本性质在线性变换和平移下保持不变. 例如，A 是一个适当阶数的矩阵，由矩阵乘法的线性性质，对 $0 \leqslant t \leqslant 1$，有

$$\begin{aligned} A\boldsymbol{x}(t) &= A[(1-t)^3 \boldsymbol{p}_0 + 3t(1-t)^2 \boldsymbol{p}_1 + 3t^2(1-t)\boldsymbol{p}_2 + t^3 \boldsymbol{p}_3] \\ &= (1-t)^3 A\boldsymbol{p}_0 + 3t(1-t)^2 A\boldsymbol{p}_1 + 3t^2(1-t)A\boldsymbol{p}_2 + t^3 A\boldsymbol{p}_3 \end{aligned}$$

新的控制点为 $A\boldsymbol{p}_0, \cdots, A\boldsymbol{p}_3$. 贝塞尔曲线平移的情况在习题 1 中考虑.

图 8-34 的曲线暗示控制点确定了曲线在起始和最终控制点处的切线. 回顾微积分中对于任何参数形式的曲线，如 $\boldsymbol{y}(t)$，曲线在点 $\boldsymbol{y}(t)$ 的切线方向由导数 $\boldsymbol{y}'(t)$ 给出，称为曲线的**切向量**.（导

数 $y'(t)$ 的计算是对 $y(t)$ 的各分量逐一求导.)

例 1 确定贝塞尔曲线 $w(t)$在 t=0 和 t=1 处的切向量是如何依赖于曲线的控制点的.

解 把方程(1)展开为多项式形式:

$$w(t)=(1-2t+t^2)p_0+(2t-2t^2)p_1+t^2p_2$$

那么由导数是函数的线性变换得,

$$w'(t)=(-2+2t)p_0+(2-4t)p_1+2tp_2$$

故

$$w'(0)=-2p_0+2p_1=2(p_1-p_0)$$
$$w'(1)=-2p_1+2p_2=2(p_2-p_1)$$

例如,在 p_0 的切向量的方向是从点 p_0 指向 p_1,大小为 p_0 到 p_1 线段长度的 2 倍. 注意,当 $p_1=p_0$ 时, $w'(0)=\mathbf{0}$. 在这种情况下, $w(t)=(1-t^2)p_1+t^2p_2$, 而且 $w(t)$ 的图形是从 p_1 到 p_2 的线段. ■

两条贝塞尔曲线的连接

两条贝塞尔曲线可以首尾相连,使得第一条 $x(t)$的终点就是第二条 $y(t)$的起点 p_2. 这样组合成的连接曲线称为(在 p_2 点)具有 G^0 几何连续性的,这是由于这两段曲线在 p_2 相连. 如果点 p_2 处曲线 1 的切线方向不同于曲线 2 的切线方向,那么一个"角"或方向的突然改变就会出现在 p_2 点. 见图 8-35.

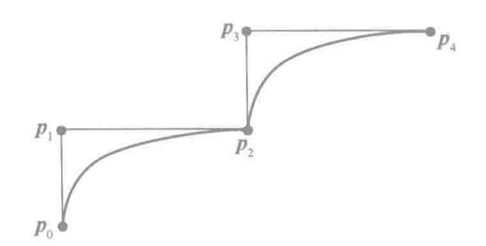


图 8-35 p_2 点处的 G^0 连续性

为了避免明显的弯曲,通常充分地调整曲线,使其具有所谓的 G^1 几何连续性,即在点 p_2 的两切向量方向相同. 导数 $x'(1)$ 与 $y'(0)$ 指向同一个方向,即使它们的大小不同. 当切向量在 p_2 相等时,切向量便在 p_2 连续,这时称连接曲线具有 C^1 连续性或 C^1 参数连续性. 图 8-36a 给出了 G^1 连续性,图 8-36b 给出了 C^1 连续性.

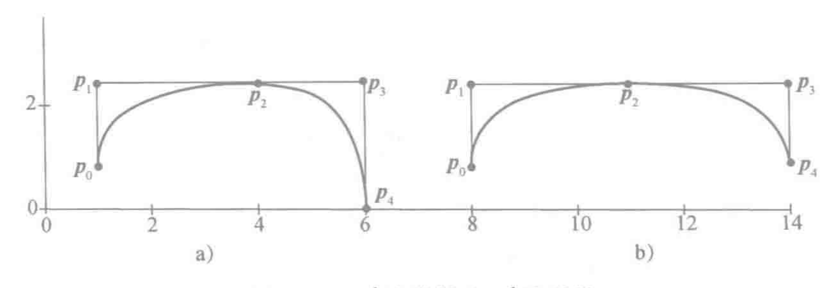


图 8-36 G^1 连续性和 C^1 连续性

例 2 设 $\boldsymbol{x}(t)$， $\boldsymbol{y}(t)$ 分别是控制点 $\{\boldsymbol{p}_0,\boldsymbol{p}_1,\boldsymbol{p}_2\}$ 与 $\{\boldsymbol{p}_2,\boldsymbol{p}_3,\boldsymbol{p}_4\}$ 所确定的两条二次贝塞尔曲线. 曲线在 $\boldsymbol{p}_2=\boldsymbol{x}(1)=\boldsymbol{y}(0)$ 相连接.

a. 设连接曲线在 $\boldsymbol{p}_2$ 具有 G^1 连续性. 此条件在控制点上产生什么代数限制？用几何语言表述此限制.

b. 若连续曲线在 $\boldsymbol{p}_2$ 具有 C^1 连续性又会怎么样？

解 a. 由例 1 知， $\boldsymbol{x}'(1)=2(\boldsymbol{p}_2-\boldsymbol{p}_1)$. 同时使用 $\boldsymbol{y}(t)$ 的控制点取代 $\boldsymbol{w}(t)$ 对应的控制点，由例 1 知， $\boldsymbol{y}'(0)=2(\boldsymbol{p}_3-\boldsymbol{p}_2)$. G^1 连续性意味着存在某个正常数 k，满足 $\boldsymbol{y}'(0)=k\boldsymbol{x}'(1)$. 等价于

$$(\boldsymbol{p}_3-\boldsymbol{p}_2)=k(\boldsymbol{p}_2-\boldsymbol{p}_1),\quad k>0 \tag{3}$$

从几何上看，(3) 式表明 $\boldsymbol{p}_2$ 在从 $\boldsymbol{p}_1$ 到 $\boldsymbol{p}_3$ 的线段上. 为了证明此结论，设 $t=(k+1)^{-1}$，注意 $0<t<1$. 解得 $k=(1-t)/t$. 把它代入 (3) 得 $\boldsymbol{p}_2=(1-t)\boldsymbol{p}_1+t\boldsymbol{p}_3$，即证明结论.

b. C^1 连续性意味着 $\boldsymbol{y}'(0)=\boldsymbol{x}'(1)$. 因而 $2(\boldsymbol{p}_3-\boldsymbol{p}_2)=2(\boldsymbol{p}_2-\boldsymbol{p}_1)$， 故 $\boldsymbol{p}_2=(\boldsymbol{p}_1+\boldsymbol{p}_3)/2$.

从几何上看， $\boldsymbol{p}_2$ 是从 $\boldsymbol{p}_1$ 到 $\boldsymbol{p}_3$ 的线段中点. 见图 8-36.

图 8-37 给出了两条三次贝塞尔曲线的 C^1 连续性. 注意，连接两段曲线之间的点在相邻控制点之间的线段的中点上.

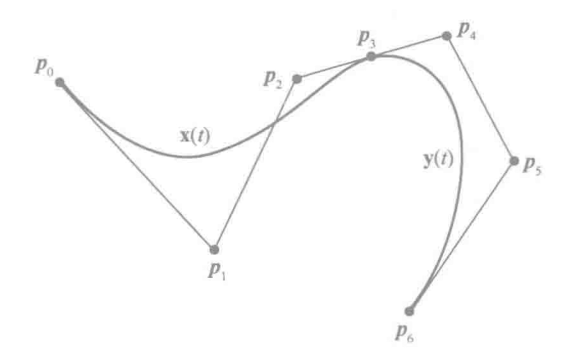


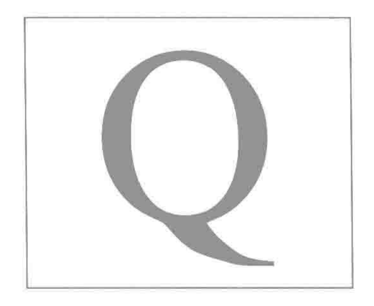
图 8-37 两条三项贝塞尔曲线

当它们具有 C^1 连续性且二阶导数 $\boldsymbol{x}''(1)$ 与 $\boldsymbol{y}''(0)$ 相等时，两曲线具有 C^2 连续性. 对于三次贝塞尔曲线这是可能的，但它很苛刻地限制控制点的位置. 另一类曲线称为 B-样条曲线，总是具 C^2 连续性，因为每对曲线共用三个控制点而不是一个. 使用 B-样条曲线绘制的图形具有更多的控制点，因而需更多的计算. 本节的一些习题能让我们更好地认识这些曲线.

奇怪的是，若 $\boldsymbol{x}(t)$， $\boldsymbol{y}(t)$ 在 $\boldsymbol{p}_3$ 相连，对于 G^1 连续性与 C^1 连续性，曲线在点 $\boldsymbol{p}_3$ 的光滑性通常是相同的. 这是由于 $\boldsymbol{x}'(t)$ 的大小与曲线的形状无关. 其大小仅反应了曲线的数学参数特性. 例如，如果一个新的向量函数 $\boldsymbol{z}(t)$ 等于 $\boldsymbol{x}(2t)$，那么 $\boldsymbol{z}(t)$ 穿过从 $\boldsymbol{p}_0$ 到 $\boldsymbol{p}_3$ 的曲线的速度是原来的 2 倍，因为当 $t=0.5$ 时，$2t$ 就到达 1. 而由微分的链式法则， $\boldsymbol{z}'(t)=2\boldsymbol{x}'(2t)$，故 $\boldsymbol{z}(t)$ 在 $\boldsymbol{p}_3$ 的切向量就是 $\boldsymbol{x}(t)$ 在 $\boldsymbol{p}_3$ 的切向量的 2 倍.

在实际应用中，通常使用许多简单的贝塞尔曲线来创建图形对象. 排版程序提供了一个重要应用的例子，因为许多字母的字体涉及弯曲线段. PostScript 字体库中的每个字母存储为一套控制点集，同时带有关于如何使用直线段与贝赛尔曲线构建字母“轮廓”的信息. 放大这样一

个字母大体需要每个控制点的坐标乘以一个常数因子. 一旦算出字母的轮廓就需要填充字母适当的实体部分. 图 8-38 说明了 PostScript 字体库中的一个字母的情形. 注意控制点.



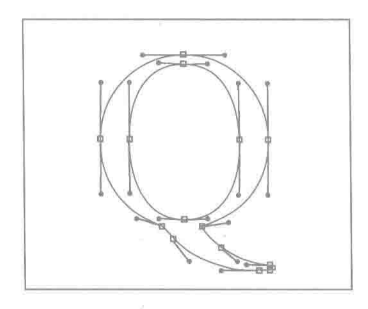


图 8-38 字体库中的一个字母

贝塞尔曲线的矩阵方程

由于贝塞尔曲线是控制点的线性组合，用多项式给出权，$\boldsymbol{x}(t)$ 的表达式可写为：

$$\boldsymbol{x}(t)=[\boldsymbol{p}_0\ \boldsymbol{p}_1\ \boldsymbol{p}_2\ \boldsymbol{p}_3]\begin{bmatrix}(1-t)^3\\3t(1-t)^2\\3t^2(1-t)\\t^3\end{bmatrix}$$

$$=[\boldsymbol{p}_0\ \boldsymbol{p}_1\ \boldsymbol{p}_2\ \boldsymbol{p}_3]\begin{bmatrix}1-3t+3t^2-t^3\\3t-6t^2+3t^3\\3t^2-3t^3\\t^3\end{bmatrix}$$

$$=[\boldsymbol{p}_0\ \boldsymbol{p}_1\ \boldsymbol{p}_2\ \boldsymbol{p}_3]\begin{bmatrix}1&-3&3&-1\\0&3&-6&3\\0&0&3&-3\\0&0&0&1\end{bmatrix}\begin{bmatrix}1\\t\\t^2\\t^3\end{bmatrix}$$

由 4 列控制点构成的矩阵称为**几何矩阵** G. 多项式系数构成的 4×4 矩阵为**贝塞尔基矩阵** M_B. 若 $\boldsymbol{u}(t)$ 为 t 的各次幂构成的列向量，则贝塞尔曲线可写成

$$\boldsymbol{x}(t)=GM_B\boldsymbol{u}(t) \tag{4}$$

在计算机图形学中，其他参数的三次曲线也用这种形式表示. 例如，若适当地改变矩阵 M_B 的元素，最后得到的曲线是 B-样条曲线. 它们比贝赛尔曲线“更光滑”，但不经过任何控制点. 当矩阵 M_B 由埃尔米特基矩阵给出时就产生**埃尔米特**三次曲线. 在此情形下，几何矩阵的列由曲线的起点与终点及曲线在这些点的切向量构成. ㊀

方程（4）中的贝塞尔曲线也可以“拆装”成另一种形式，以用于讨论贝塞尔曲面. 为了之

㊀ 基矩阵是由用于定义曲线的混合多项式系数组成的矩阵的行构成的. 对于三次贝塞尔曲线，4 个多项式为$(1-t)^3$，$3t(1-t)^2$，$3t^2(1-t)$和 t^3. 它们构成了 3 阶及以下多项式空间 $\mathbb{P}_3$ 的基. 向量 $\boldsymbol{x}(t)$中每一项都是这些多项式的线性组合. 权重是（4）中几何矩阵 G 的行.

后叙述的方便，参数 t 由参数 s 代替：

$$\boldsymbol{x}(s)=\boldsymbol{u}(s)^{\mathrm{T}}M_B^{\mathrm{T}}\begin{bmatrix}\boldsymbol{p}_0\\\boldsymbol{p}_1\\\boldsymbol{p}_2\\\boldsymbol{p}_3\end{bmatrix}=\begin{bmatrix}1 & s & s^2 & s^3\end{bmatrix}\begin{bmatrix}1 & 0 & 0 & 0\\-3 & 3 & 0 & 0\\3 & -6 & 3 & 0\\-1 & 3 & -3 & 1\end{bmatrix}\begin{bmatrix}\boldsymbol{p}_0\\\boldsymbol{p}_1\\\boldsymbol{p}_2\\\boldsymbol{p}_3\end{bmatrix}$$
$$=\begin{bmatrix}(1-s)^3 & 3s(1-s)^2 & 3s^2(1-s) & s^3\end{bmatrix}\begin{bmatrix}\boldsymbol{p}_0\\\boldsymbol{p}_1\\\boldsymbol{p}_2\\\boldsymbol{p}_3\end{bmatrix} \tag{5}$$

这个公式与（4）式右边乘积的转置并不相同，因为在（5）式中 $\boldsymbol{x}(s)$和控制点没有用转置符号.（5）式中控制点构成的矩阵称为**几何向量**. 这个几何向量视为 4×1 块（分块）矩阵，它们的元素是列向量.（5）式的第二部分中，几何向量左边的矩阵也视为分块矩阵，每块是一个标量. 分块矩阵乘法是有意义的，因为几何向量中的每个元素（向量）不仅可以左乘一个矩阵也可以乘以一个标量. 因而，列向量 $\boldsymbol{x}(s)$由（5）式给出了表示.

贝塞尔曲面

4 条贝塞尔曲线的集合可构造一个三维双三次曲面片. 考虑 4 个几何矩阵

$$\begin{bmatrix}\boldsymbol{p}_{11} & \boldsymbol{p}_{12} & \boldsymbol{p}_{13} & \boldsymbol{p}_{14}\end{bmatrix}$$
$$\begin{bmatrix}\boldsymbol{p}_{21} & \boldsymbol{p}_{22} & \boldsymbol{p}_{23} & \boldsymbol{p}_{24}\end{bmatrix}$$
$$\begin{bmatrix}\boldsymbol{p}_{31} & \boldsymbol{p}_{32} & \boldsymbol{p}_{33} & \boldsymbol{p}_{34}\end{bmatrix}$$
$$\begin{bmatrix}\boldsymbol{p}_{41} & \boldsymbol{p}_{42} & \boldsymbol{p}_{43} & \boldsymbol{p}_{44}\end{bmatrix}$$

回顾方程（4），当这些矩阵右乘如下权重的向量就产生一条贝塞尔曲线：

$$M_B\boldsymbol{u}(t)=\begin{bmatrix}(1-t)^3\\3t(1-t)^2\\3t^2(1-t)\\t^3\end{bmatrix}$$

设 G 是 4×4 分块矩阵，它们的元素是以上的控制点 $\boldsymbol{p}_{ij}$. 那么下面的积就是 4×1 分块矩阵，且每个元素就是一条贝塞尔曲线：

$$GM_B\boldsymbol{u}(t)=\begin{bmatrix}\boldsymbol{p}_{11} & \boldsymbol{p}_{12} & \boldsymbol{p}_{13} & \boldsymbol{p}_{14}\\\boldsymbol{p}_{21} & \boldsymbol{p}_{22} & \boldsymbol{p}_{23} & \boldsymbol{p}_{24}\\\boldsymbol{p}_{31} & \boldsymbol{p}_{32} & \boldsymbol{p}_{33} & \boldsymbol{p}_{34}\\\boldsymbol{p}_{41} & \boldsymbol{p}_{42} & \boldsymbol{p}_{43} & \boldsymbol{p}_{44}\end{bmatrix}\begin{bmatrix}(1-t)^3\\3t(1-t)^2\\3t^2(1-t)\\t^3\end{bmatrix}$$

事实上，

$$GM_B\boldsymbol{u}(t)=\begin{bmatrix}(1-t)^3\boldsymbol{p}_{11}+3t(1-t)^2\boldsymbol{p}_{12}+3t^2(1-t)\boldsymbol{p}_{13}+t^3\boldsymbol{p}_{14}\\(1-t)^3\boldsymbol{p}_{21}+3t(1-t)^2\boldsymbol{p}_{22}+3t^2(1-t)\boldsymbol{p}_{23}+t^3\boldsymbol{p}_{24}\\(1-t)^3\boldsymbol{p}_{31}+3t(1-t)^2\boldsymbol{p}_{32}+3t^2(1-t)\boldsymbol{p}_{33}+t^3\boldsymbol{p}_{34}\\(1-t)^3\boldsymbol{p}_{41}+3t(1-t)^2\boldsymbol{p}_{42}+3t^2(1-t)\boldsymbol{p}_{43}+t^3\boldsymbol{p}_{44}\end{bmatrix}$$

现在固定 t，那么 $GM_B\boldsymbol{u}(t)$ 是列向量，它可以作为方程（5）（即以 s 为参数的贝塞尔曲线）中的几何向量. 这样就产生了**贝塞尔双三次曲面**：

$$\boldsymbol{x}(s,t)=\boldsymbol{u}(s)^{\mathrm{T}}M_B^{\mathrm{T}}\,GM_B\boldsymbol{u}(t)\text{，这里 }0\leqslant s,t\leqslant 1 \tag{6}$$

$\boldsymbol{x}(s,t)$ 的公式是 16 个控制点的线性组合. 如果想象这些控制点被布成了比较均匀的正方形排列，如图 8-39 所示，于是一个网状的 8 条贝塞尔曲线就控制并构造出了一个贝塞尔曲面，4 条沿 s 方向，还有 4 条沿 t 方向. 这个曲面总是经过“角”的 4 个控制点. 当这个“16 点曲面”是更大的一个曲面的一部分时，它的 12 个边界控制点是与相邻“16 点曲面”的边界控制点重合的.

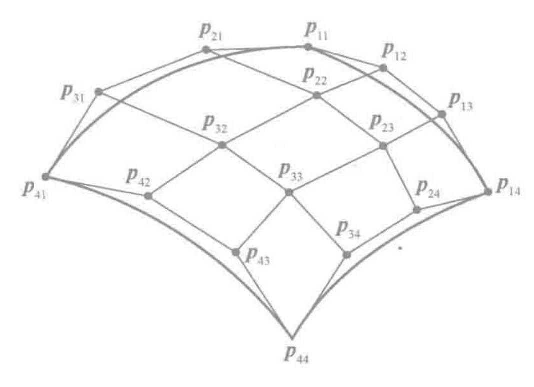


图 8-39　贝塞尔双三次曲面片的 16 个控制点

曲线与曲面的逼近

在 CAD 程序和创建虚拟现实的计算机游戏程序中，设计者通常在图形工作站中创建涉及各种几何结构的场景. 构造场景中的形体对象的过程需要设计者与对象互动. 对象的每个小变动都需要图形程序进行新的数学计算. 贝塞尔曲线与曲面在这个过程中起着非常关键的作用，各个对象需要在较少的控制点条件下，用尽可能多的多边形近似逼近. 这奇迹般地减少了计算时间，从而提高了设计者的工作效率.

然而，在场景形体对象创造以后，最终图像的准备工作常出现难易不同的计算要求. 遇到由平坦的面与笔直的棱构成的对象（如多面体）会容易计算得多.

例如，设计者要润色情境，这需要引进光源，给曲面着色加纹理，模拟曲面的反射等. 为此需要作必要的计算. 计算曲面在点 $\boldsymbol{p}$ 上反射光的方向，需要知道入射光与曲面法向量，即垂直于曲面 $\boldsymbol{p}$ 点切面的正交向量的方向. 要计算这些法向量，在由细小平坦的多边形构成的表面上要比在法向量随点 $\boldsymbol{p}$ 移动而连续变化的曲面上容易得多. 如果 $\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3$ 是一个平坦多边形的相邻顶点，那么曲面的法向量的算法很简单，是一个叉积或其反向，叉积为 $(\boldsymbol{p}_2-\boldsymbol{p}_1)\times(\boldsymbol{p}_3-\boldsymbol{p}_2)$. 当多边形很小时，只需要一个法向量来润色整个多边形. 同样，两种广泛使用的阴影程序——高氏阴影与冯氏着色，也需要由多边形来定义表面.

对平坦表面的这些需求，导致那些在场景创建平台上构造出来的贝塞尔曲线和曲面常常要用直线段和多边形平面去逼近. 逼近贝塞尔曲线或曲面的基本思想是用越来越多的控制点把曲线或曲面分成更小的曲面片.

贝塞尔曲线与曲面的递归细分

图 8-40 给出了一条贝塞尔曲线的 4 个控制点 $\boldsymbol{p}_0,\cdots,\boldsymbol{p}_3$ 以及两条新曲线的控制点，每条与原曲线的一半重叠．“左”边的一条以 $\boldsymbol{q}_0=\boldsymbol{p}_0$ 为起点，以原曲线的中点 $\boldsymbol{q}_3$ 为终点．“右”边的一条以 $\boldsymbol{r}_0=\boldsymbol{q}_3$ 为起点，以 $\boldsymbol{r}_3=\boldsymbol{p}_3$ 为终点．

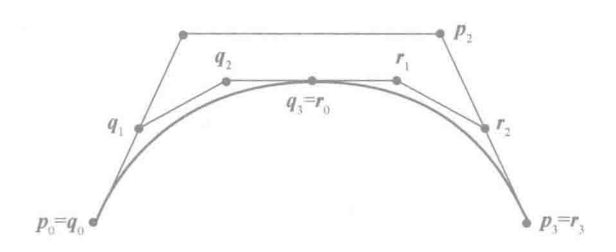


图 8-40　贝塞尔曲线的细分

图 8-41 表明新的控制点围住的部分比原来的控制点围住的部分更“瘦小”．随着控制点间的距离的减小，每条曲线段的控制点也更接近于一条直线段．贝塞尔曲线这种变量-减小性质与贝塞尔曲线总是在控制点的凸包中的事实相关．

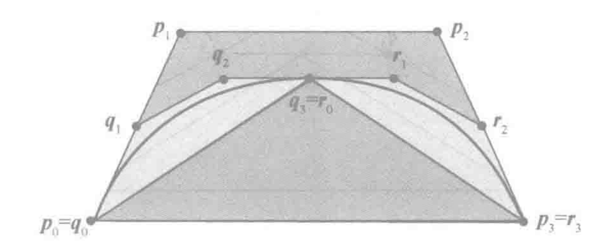


图 8-41　控制点的凸包

可以用简单的公式给出新控制点与原控制点的关系．当然，$\boldsymbol{q}_0=\boldsymbol{p}_0$ 且 $\boldsymbol{r}_3=\boldsymbol{p}_3$．当原曲线 $\boldsymbol{x}(t)$ 有标准参数表达式时，曲线 $\boldsymbol{x}(t)$ 的中点就是 $\boldsymbol{x}(0.5)$．

$$\boldsymbol{x}(t)=(1-3t+3t^2-t^3)\boldsymbol{p}_0+(3t-6t^2+3t^3)\boldsymbol{p}_1+(3t^2-3t^3)\boldsymbol{p}_2+t^3\boldsymbol{p}_3 \qquad (7)$$

其中 $0\leqslant t\leqslant 1$．因而，新的控制点 $\boldsymbol{r}_0$ 与 $\boldsymbol{q}_3$ 由下式给出：

$$\boldsymbol{q}_3=\boldsymbol{r}_0=\boldsymbol{x}(0.5)=\frac{1}{8}(\boldsymbol{p}_0+3\boldsymbol{p}_1+3\boldsymbol{p}_2+\boldsymbol{p}_3) \qquad (8)$$

剩下的“内部”控制点的公式也是简单的，但是公式的推导需要一些有关曲线的切向量的工作．由定义，参数表达形式的曲线 $\boldsymbol{x}(t)$ 的切向量为 $\boldsymbol{x}'(t)$．这个向量给出了曲线在 $\boldsymbol{x}(t)$ 点的切线的方向．对于（7）式的贝塞尔曲线，有

$$\boldsymbol{x}'(t)=(-3+6t-3t^2)\boldsymbol{p}_0+(3-12t+9t^2)\boldsymbol{p}_1+(6t-9t^2)\boldsymbol{p}_2+3t^2\boldsymbol{p}_3$$

其中 $0\leqslant t\leqslant 1$．特别地，

$$\boldsymbol{x}'(0)=3(\boldsymbol{p}_1-\boldsymbol{p}_0)\text{ 且 }\boldsymbol{x}'(1)=3(\boldsymbol{p}_3-\boldsymbol{p}_2) \qquad (9)$$

从几何上看，$\boldsymbol{p}_1$ 在曲线于点 $\boldsymbol{p}_0$ 的切线上，而 $\boldsymbol{p}_2$ 在曲线于点 $\boldsymbol{p}_3$ 的切线上，见图 8-41．同时由 $\boldsymbol{x}'(t)$，计算得，

$$x'(0.5) = \frac{3}{4}(-\boldsymbol{p}_0 - \boldsymbol{p}_1 + \boldsymbol{p}_2 + \boldsymbol{p}_3) \tag{10}$$

设 $\boldsymbol{y}(t)$ 为由 $\boldsymbol{q}_0,\cdots,\boldsymbol{q}_3$ 确定的贝塞尔曲线且 $\boldsymbol{z}(t)$ 是由 $\boldsymbol{r}_0,\cdots,\boldsymbol{r}_3$ 确定的贝塞尔曲线. 由于 $\boldsymbol{y}(t)$ 与 $\boldsymbol{x}(t)$ 穿过同样的路径，但 t 从 0 到 1 变化时达到 $\boldsymbol{x}(0.5), \boldsymbol{y}(t) = \boldsymbol{x}(0.5t)$，$0 \leqslant t \leqslant 1$. 类似地，由于 $\boldsymbol{z}(t)$ 当 t=0 时以 $\boldsymbol{x}(0.5)$为起点，$\boldsymbol{z}(t) = \boldsymbol{x}(0.5 + 0.5t), 0 \leqslant t \leqslant 1$. 由导数的链式法则得

$$\boldsymbol{y}'(t) = 0.5\boldsymbol{x}'(0.5t) \text{ 且 } \boldsymbol{z}'(t) = 0.5\boldsymbol{x}'(0.5 + 0.5t), \quad 0 \leqslant t \leqslant 1 \tag{11}$$

在（9）式中用 $\boldsymbol{y}'(0)$ 替代 $\boldsymbol{x}'(0)$，（11）式中令 t=0，由（9）式，$\boldsymbol{y}(t)$ 的控制点满足

$$3(\boldsymbol{q}_1 - \boldsymbol{q}_0) = \boldsymbol{y}'(0) = 0.5\boldsymbol{x}'(0) = \frac{3}{2}(\boldsymbol{p}_1 - \boldsymbol{p}_0) \tag{12}$$

在（9）式中用 $\boldsymbol{y}'(1)$ 替代 $\boldsymbol{x}'(1)$，（11）式中令 t=1，由（10）式，

$$3(\boldsymbol{q}_3 - \boldsymbol{q}_2) = \boldsymbol{y}'(1) = 0.5\boldsymbol{x}'(0.5) = \frac{3}{8}(-\boldsymbol{p}_0 - \boldsymbol{p}_1 + \boldsymbol{p}_2 + \boldsymbol{p}_3) \tag{13}$$

由习题 13 所示，解方程（8），（9），（10），（12）和（13）可得到 $\boldsymbol{q}_0,\cdots,\boldsymbol{q}_3$ 的表达式. 图 8-42 给出了公式的几何意义. 内部控制点 $\boldsymbol{q}_1$，$\boldsymbol{r}_2$ 分别是线段 $\boldsymbol{p}_0\boldsymbol{p}_1$ 与线段 $\boldsymbol{p}_2\boldsymbol{p}_3$ 的中点. 把线段 $\boldsymbol{p}_1\boldsymbol{p}_2$ 的中点与 $\boldsymbol{q}_1$ 连接起来，得到的线段以 $\boldsymbol{q}_2$ 为中点！

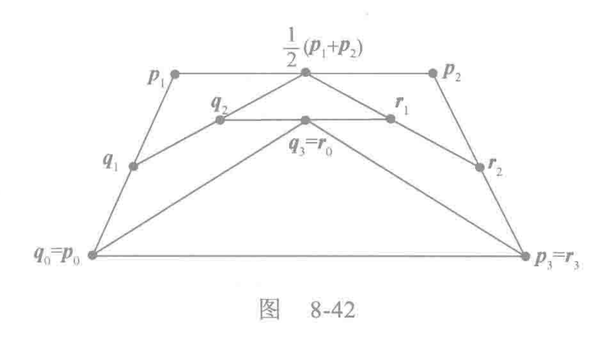


图 8-42

以上完成了细分过程的第一步. 现在开始递归，细分两条新的曲线. 递归进行到所有的曲线都充分地直为止. 每步的递归都是可以选择地进行的，即若两条新曲线中的一条已经充分地直，则可以不细分这一条，只需细分另一条. 一旦细分完全停止，每条曲线的两端点之间用直线段连接，如此继续下去，最终图像准备工作的每一步即告完成.

贝塞尔双三次曲面也与贝塞尔曲线一样具有变量-减小性质，因此对于曲面上每一条交叉的贝塞尔曲线都可以使用上述过程. 略去细节，这里只提一下基本思路. 考虑 4 条“平行”贝塞尔曲线，其参数是 s，对它们进行一次细分，每条上的控制点都增加到同一数量 8，每条上的 8 个控制点确定了参数 s 从 0 到 1 的曲线，共 4 条. 然后，对 t 变量，便有了 8 条贝塞尔曲线要处理，每条有 4 个控制点，对它们进行一次递归细分，每条上的控制点也都增加到同一数量 8. 于是，总共得到了 64 个控制点. 对曲面递归细分的每一步也与前面对贝塞尔曲线那样是可以选择地进行的，但需要面对一些困难的细节问题.[㊀]

㊀ 参见 Foley，van Dam, Feiner, and Hughes, *Computer Graphics-Principles and Practice*, 2nd Ed. (Boston:Addison-Wesley，1996)，pp. 527-528.

练习题

样条曲线通常是关于过特殊点的曲线. 然而 B-样条曲线通常不过它的控制点. 一条 B-样条曲线的参数形式为:

$$x(t)=\frac{1}{6}\left[(1-t)^3 p_0+(4-6t^2+3t^3)p_1+(1+3t+3t^2-3t^3)p_2+t^3 p_3\right] \tag{14}$$

其中 $0\leqslant t\leqslant 1$，而 p_0， p_1， p_2 和 p_3 是控制点. 当 t 从 0 到 1 时， $x(t)$ 创建了一条简短的临近 $\overline{p_1p_2}$ 的曲线. 基本的代数知识表明上述 B-样条曲线的公式也可以写为

$$x(t)=\frac{1}{6}\left[(1-t)^3 p_0+(4-3t+3t(1-t)^2)p_1+(1+3t+3t^2(1-t))p_2+t^3 p_3\right] \tag{15}$$

这表明了它与贝塞尔曲线的类似性. 除了前面的因子 1/6， p_0 与 p_3 的项相同. p_1 项增加了 $4-3t$， p_2 项增加了 $1+3t$. 较之贝塞尔曲线，这些项使得曲线更加接近 $\overline{p_1p_2}$. 为了保证系数和为 1，因子 1/6 是必要的. 图 8-43 对比了具有相同控制点的 B-样条线段与贝塞尔曲线.

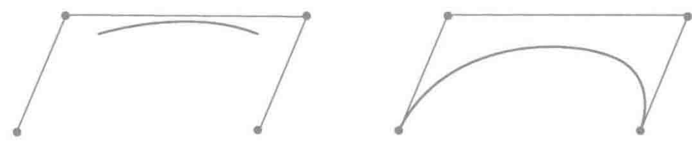


图 8-43　B-样条线段和贝塞尔曲线

1. 证明：B-样条曲线不是以 p_0 为起点，但是 $x(0)$ 在 $\text{conv}\{p_0,p_1,p_2\}$ 里. 假设 p_0， p_1， p_2 是仿射无关的，求 $x(0)$ 对应于 $\{p_0,p_1,p_2\}$ 的仿射坐标.

2. 证明：B-样条曲线不是以 p_3 为起点，但是 $x(1)$ 在 $\text{conv}\{p_1,p_2,p_3\}$ 里. 假设 p_1， p_2， p_3 是仿射无关的，求 $x(1)$ 对应于 $\{p_1,p_2,p_3\}$ 的仿射坐标.

习题 8.6

1. 设贝塞尔曲线平移为 $x(t)+b$. 即对于 $0\leqslant t\leqslant 1$，新曲线为：
   $x(t)=(1-t)^3 p_0+3t(1-t)^2 p_1+3t^2(1-t)p_2+t^3 p_3+b$
   证明：新曲线还是一条贝塞尔曲线.（提示：新的控制点在哪里?）
2. 练习题中所定义的 B-样条曲线的参数向量表达式为：
   $$x(t)=\frac{1}{6}\left[(1-t)^3 p_0+(4-3t+3t(1-t))p_1+(1+3t+3t^2(1-t))p_2+t^3 p_3\right],\quad 0\leqslant t\leqslant 1$$
   这里 p_0， p_1， p_2， p_3 是控制点.
   a. 证明：对 $0\leqslant t\leqslant 1$， $x(t)$ 在控制点的凸包中.
   b. 假设 B-样条曲线 $x(t)$ 平移为 $x(t)+b$（类似于习题 1）. 证明：新的曲线也是 B-样条曲线.
3. 设 $x(t)$ 是由点 p_0， p_1， p_2 和 p_3 确定的三次贝塞尔曲线.
   a. 计算切向量 $x'(t)$. 确定 $x'(0)$ 与 $x'(1)$ 同控制点的关系，并给出这些切向量方向的几何描述. $x'(1)=\mathbf{0}$ 可能吗？
   b. 计算二次导数 $x''(t)$ 并确定 $x''(0)$ 与 $x''(1)$ 同控制点的关系. 画出基于图 8-43 的图形，并构造一段指向 $x''(0)$ 方向的线段.（提示：把 p_1 视为坐标系的原点.）
4. 设 $x(t)$ 是习题 2 中的 B-样条曲线，以 p_0， p_1， p_2 和 p_3 为控制点.
   a. 计算切向量 $x'(t)$ 并确定 $x'(0)$ 与 $x'(1)$ 同控制点的关系. 给出这些切向量方向的几何描述. 思考当 $x'(1)$ 与 $x'(0)$ 都等于 $\mathbf{0}$ 会怎样？证明你的结论.

b. 计算二次导数 $x''(t)$ 并确定 $x''(0)$ 与 $x''(1)$ 同控制点的关系. 画出基于图 8-43 的图形，并构造一段指向 $x''(1)$ 方向的线段. （提示：把 p_2 视为坐标系的原点. ）

5. 设 $x(t)$ ， $y(t)$ 是分别由控制点 $\{p_0, p_1, p_2, p_3\}$ 与 $\{p_3, p_4, p_5, p_6\}$ 确定的三次贝塞尔曲线，且 $x(t)$ 和 $y(t)$ 在 p_3 相连接. 下面的问题是关于由 $x(t)$ 连接 $y(t)$ 组成的曲线的. 为简单起见，假定曲线是 $\mathbb{R}^2$ 空间的.

a. 控制点具有什么条件将保证曲线在 p_3 具有 C^1 连续性. 证明你的答案.

b. $x'(1)$ 与 $y'(0)$ 都等于 $\mathbf{0}$ 向量会怎样？

6. 设B-样条曲线由习题 2 中所描述的B-样条线段构成. $p_0,\cdots,p_4$ 是控制点，对 $0 \leqslant t \leqslant 1$，设 $x(t)$ ， $y(t)$ 分别由几何矩阵 $[p_0\ \ p_1\ \ p_2\ \ p_3]$ 与 $[p_1\ \ p_2\ \ p_3\ \ p_4]$ 确定. 注意两线段共享三个控制点. 两线段并不重叠，但它们在临近 p_2 的一个共同端点连接.

a. 证明：连接曲线具有 G^0 连续性，即 $x(1) = y(0)$.

b. 证明：曲线在连接点 $x(1)$ 具有 C^1 连续性，即证明 $x'(1) = y'(0)$.

7. 设 $x(t)$ ， $y(t)$ 是习题 5 中的贝塞尔曲线，假定连接曲线在 p_3 具有 C^2 连续性(包含 C^1 连续性). 令 $x''(1) = y''(0)$ ，证明： p_5 完全由 p_1 ， p_2 和 p_3 确定. 因而点 $p_0,\cdots,p_3$ 与 C^2 条件确定所有的控制点，除了 $y(t)$ 的一个控制点.

8. 设 $x(t)$ ， $y(t)$ 是习题 6 中的 B-样条曲线的部分. 证明：曲线在点 $x(1)$ 具有 C^2 连续性（包含 C^1 连续性，即证明 $x''(1) = y''(0)$ ）. 这种高阶连续性在 CAD 应用（如汽车车身设计）中是令人满意的，因为曲线与曲面表现得相当光滑. 然而，对于长度可比的曲线，B-样条曲线的计算量是贝塞尔曲线的 3 倍. 对于曲面，B-样条曲线的计算量是贝塞尔曲面的 9 倍. 程序设计者在需要实时表现的应用（如飞机座舱模拟）中会选择贝塞尔曲面.

9. 由 5 个控制点 p_0 ， p_1 ， p_2 ， p_3 和 p_4 确定的二次贝塞尔曲线如下：

$$x(t) = (1-t)^4 p_0 + 4t(1-t)^3 p_1 + 6t^2(1-t)^2 p_2 + 4t^3(1-t)p_3 + t^4 p_4 , \quad 0 \leqslant t \leqslant 1$$

构造 $x(t)$ 的二次基矩阵 M_B .

10. B-样条曲线中的“B”是基于片段 $x(t)$ 可以写成基矩阵 M_S 的形式，类似于贝塞尔曲线的公式，即

$$x(t) = GM_S u(t) , \quad 0 \leqslant t \leqslant 1$$

这里 G 是几何矩阵 $[p_0\ \ p_1\ \ p_2\ \ p_3]$ ， $u(t)$ 是列向量 $(1, t, t^2, t^3)$. 对于均匀的 B-样条曲线，每段使用相同的基矩阵，但几何矩阵是变化的. 构造 $x(t)$ 的基矩阵 M_S .

在习题 11 和 12 中，判断正误，并证明.

11. a. 三次贝塞尔曲线基于 4 个控制点.

b. 给定以 p_0, p_1, p_2 为控制点的二次贝塞尔曲线 $x(t)$ ，方向线段 $p_1 - p_0$ （从 p_0 到 p_1 ）是曲线在 p_0 的切向量.

c. 当两条以 $\{p_0, p_1, p_2\}$ 与 $\{p_2, p_3, p_4\}$ 为控制点的二次贝塞尔曲线在 p_2 相连接. 若 p_2 是线段 p_1 到 p_3 的中点，则连接的贝塞尔曲线在 p_2 具有 C^1 连续性.

12. a. 在线性变换下，贝塞尔曲线的基本性质保持不变. 但在平移下不是如此.

b. 当两条贝塞尔曲线 $x(t)$ ， $y(t)$ 在点 $x(1) = y(0)$ 连接，则连接曲线在此点具有 G^0 连续性.

c. 贝塞尔基矩阵是一个以曲线的控制点为列的矩阵.

习题 13~15 是有关图 8-40 所示的贝塞尔曲线的细分问题. 设 $x(t)$ 以 $p_0,\cdots,p_3$ 为控制点且 $y(t)$ 与 $z(t)$ 是细分的贝塞尔曲线，分别以 $q_0,\cdots,q_3$ 与 $r_0,\cdots,r_3$ 为控制点.

13. a. 使用方程（12）证明： q_1 是从 p_0 到 p_1 的线段的中点.

b. 使用方程（13）证明：

$$8\boldsymbol{q}_2 = 8\boldsymbol{q}_3 + \boldsymbol{p}_0 + \boldsymbol{p}_1 - \boldsymbol{p}_2 - \boldsymbol{p}_3$$

c. 使用（b）、方程（8）及（a）证明：$\boldsymbol{q}_2$ 是从 $\boldsymbol{q}_1$ 到 $\boldsymbol{p}_0$ 与 $\boldsymbol{p}_1$ 的线段中点的线段中点. 即

$$\boldsymbol{q}_2 = \frac{1}{2}\left[\boldsymbol{q}_1 + \frac{1}{2}(\boldsymbol{p}_1 + \boldsymbol{p}_2)\right]$$

14. a. 验证每个等号：

$$3(\boldsymbol{r}_3 - \boldsymbol{r}_2) = \boldsymbol{z}'(1) = 0.5\boldsymbol{x}'(1) = \frac{3}{2}(\boldsymbol{p}_3 - \boldsymbol{p}_2)$$

b. 证明 $\boldsymbol{r}_2$ 是 $\boldsymbol{p}_3$ 到 $\boldsymbol{p}_2$ 线段的中点.

c. 验证每个等号：$3(\boldsymbol{r}_1 - \boldsymbol{r}_0) = \boldsymbol{z}'(0) = 0.5\boldsymbol{x}'(0.5)$

d. 使用（c）证明：$8\boldsymbol{r}_1 = -\boldsymbol{p}_0 - \boldsymbol{p}_1 + \boldsymbol{p}_2 + \boldsymbol{p}_3 + 8\boldsymbol{r}_0$.

e. 使用（d）、（8）式及（a）证明：$\boldsymbol{r}_1$ 是 $\boldsymbol{r}_2$ 到 $\boldsymbol{p}_1$ 与 $\boldsymbol{p}_2$ 的线段中点的线段中点，即

$$\boldsymbol{r}_1 = \frac{1}{2}[\boldsymbol{r}_2 + \frac{1}{2}(\boldsymbol{p}_1 + \boldsymbol{p}_2)].$$

15. 有时候一条贝塞尔曲线只有一半需要进一步细分，这种情况下，例如，要细分“左”边，用习题 13 中（a）和（c）及（8）式可以完成. 而当曲线的两半边都需要细分时，可以设计一个计算过程直接有效地协同确定左右两半边的控制点，使之不用（8）式.

a. 证明：切向量 $\boldsymbol{y}'(1)$ 和 $\boldsymbol{z}'(0)$ 是相等的.

b. 用（a）证明：$\boldsymbol{q}_3$（它等于 $\boldsymbol{r}_0$）是从 $\boldsymbol{q}_2$ 到 $\boldsymbol{r}_1$ 的线段的中点.

c. 用（b）及习题 13 和 14 写一个计算 $\boldsymbol{y}(t)$ 和 $\boldsymbol{z}(t)$ 控制点的高效算法. 涉及的运算只有求和及除以 2.

16. 解释为什么三次贝塞尔曲线完全由 $\boldsymbol{x}(0)$，$\boldsymbol{x}'(0)$，$\boldsymbol{x}(1)$ 和 $\boldsymbol{x}'(1)$ 确定.

17. 由苹果公司与 Adobe 系统创立的 TrueType 字体使用二次贝塞尔曲线，而由微软公司创立的 PostScript 字体使用三次贝塞尔曲线. 三次曲线为字体的设计提供更大的灵活性，但把使用二次曲线的每个字体转化为使用三次曲线对微软公司是重要的. 假定 $\boldsymbol{w}(t)$ 是以 $\boldsymbol{p}_0$，$\boldsymbol{p}_1$ 和 $\boldsymbol{p}_2$ 为控制点的二次曲线.

a. 求点 $\boldsymbol{r}_0, \boldsymbol{r}_1, \boldsymbol{r}_2$ 和 $\boldsymbol{r}_3$，使得以这些点为控制点的三次贝塞尔曲线 $\boldsymbol{x}(t)$ 与 $\boldsymbol{w}(t)$ 具有相同的起点与终点且在 $t=0$ 与 $t=1$ 有相同的切向量.（见习题 16.）

b. 证明：若 $\boldsymbol{x}(t)$ 如(a)一样构造，则 $\boldsymbol{x}(t) = \boldsymbol{w}(t)$，$0 \leqslant t \leqslant 1$.

18. 使用分块矩阵乘法计算下面的矩阵乘积，这在贝塞尔曲线的另一公式（5）中出现：

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 3 & 0 & 0 \\ 3 & -6 & 0 & 0 \\ -1 & 3 & -3 & 1 \end{bmatrix}\begin{bmatrix} \boldsymbol{p}_0 \\ \boldsymbol{p}_1 \\ \boldsymbol{p}_2 \\ \boldsymbol{p}_3 \end{bmatrix}$$

练习题答案

1. 由方程（14）有 $t=0, \boldsymbol{x}(0) \neq \boldsymbol{p}_0$，因为

$$\boldsymbol{x}(0) = \frac{1}{6}[\boldsymbol{p}_0 + 4\boldsymbol{p}_1 + \boldsymbol{p}_2] = \frac{1}{6}\boldsymbol{p}_0 + \frac{2}{3}\boldsymbol{p}_1 + \frac{1}{6}\boldsymbol{p}_2$$

系数是非负的且和为 1，故 $\boldsymbol{x}(0)$ 在 $\text{conv}\{\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2\}$ 里，且对应于 $\{\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2\}$ 的仿射坐标是 $\left(\frac{1}{6}, \frac{2}{3}, \frac{1}{6}\right)$.

2. 由方程（14）有 $t=1, \boldsymbol{x}(1) \neq \boldsymbol{p}_3$，因为

$$\boldsymbol{x}(1) = \frac{1}{6}[\boldsymbol{p}_1 + 4\boldsymbol{p}_2 + \boldsymbol{p}_3] = \frac{1}{6}\boldsymbol{p}_1 + \frac{2}{3}\boldsymbol{p}_2 + \frac{1}{6}\boldsymbol{p}_3$$

系数是非负的且和为 1，故 $\boldsymbol{x}(1)$ 在 $\text{conv}\{\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3\}$ 里，且对应于 $\{\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3\}$ 的仿射坐标是 $\left(\frac{1}{6}, \frac{2}{3}, \frac{1}{6}\right)$.

附录 A　简化形阶梯矩阵的唯一性

定理（简化形阶梯矩阵的唯一性）

每一个 $m\times n$ 矩阵 A 行等价于唯一简化形梯形矩阵 U.

证　利用 4.3 节的思想，行等价矩阵的列具有完全一样的线性相关性.

行推导的算法说明至少存在一个这样的矩阵 U. 假设 A 行等价于简化形式的阶梯矩阵 U 和 V，U 的行中最左边非零元素是“主元”1，称这类主元 1 的位置是一个主元位置，并且包含它的列称为主元列，（这个定义仅用于阶梯特征的 U 和 V，并且假设简化形式的阶梯形不唯一）.

U 和 V 的主元列恰恰是与它们左边的列线性无关的非零列，（这个条件自动满足第一列是非零的）. 由于 U 和 V 行等价，（两个矩阵都行等价于 A），它们的列有相同的线性相关性，因此，U 和 V 的主元列出现在同样位置，如果有 r 个这样的列，又因为 U 和 V 是简化形式的阶梯形式，它们的主元列是 $m\times n$ 单位矩阵的前 r 列，这样对应 U 和 V 的主元列是相等的.

最后，考虑 U 的任意非主元列，例如列 j，这个列或者是零或者是左边主元列的线性组合（因为这些主元列是第 j 列左边列生成空间的一个基）. 两种情形下，对第 j 个元素为 1 的 x 都可表示写成 $U\mathbf{x}=\mathbf{0}$，那么也有 $V\mathbf{x}=\mathbf{0}$，这说明 V 的第 j 列或者是零或者同样是它左边 V 的主元列的线性组合，由于 U 和 V 对应的主元列是相等的，U 和 V 的第 j 列也相等，这个结果对 U 和 V 所有非主元列相等，这就证明了 U 是唯一的. ■

附录B 复　　数

复数是可以写成如下形式的数：

$$z = a + bi$$

其中 a 和 b 是实数，i 是满足关系 $i^2 = -1$ 的常用符号，数 a 是 z 的**实部**，记作 $\operatorname{Re} z$，数 b 是 z 的**虚部**，记作 $\operatorname{Im} z$，两个复数相等当且仅当它们的实部和虚部分别相等．例如，如果 $z = 5 + (-2)i$，那么 $\operatorname{Re} z = 5$ 且 $\operatorname{Im} z = -2$，可以简记为 $z = 5 - 2i$．

一个实数 a 可以认为是一个特殊类型的复数，将 a 和 $a + 0i$ 作为一个数．更进一步，实数上的算术运算可以扩展到复数集合上．

复数系是所有复数的集合，记作 $\mathbb{C}$，并具有下面加法和标量乘法运算：

$$(a + bi) + (c + di) = (a + c) + (b + d)i \tag{1}$$

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i \tag{2}$$

这些法则当（1）和（2）式中 b 和 d 为零时，是普通实数的加法和乘法．很容易验证，在 $\mathbb{R}$ 上的常见法则在 $\mathbb{C}$ 上同样成立．由于这个原因，乘法常用扩展的代数法则来计算．

例 1

$$\begin{aligned}(5 - 2i)(3 + 4i) &= 15 + 20i - 6i - 8i^2 \\ &= 15 + 14i - 8(-1) \\ &= 23 + 14i\end{aligned}$$

也就是说，用项 $(3 + 4i)$ 乘 $5 - 2i$ 的每一项，利用 $i^2 = -1$，并将结果写成 $a + bi$ 的形式．■

复数 z_1 和 z_2 的减法定义为：

$$z_1 - z_2 = z_1 + (-1)z_2$$

特别地，我们用 $-z$ 代替 $(-1)z$．

复数 $z = a + bi$ 的共轭复数是复数 $\bar{z}$（读作"z 棒"），定义为

$$\bar{z} = a - bi$$

$\bar{z}$ 可用 z 中相反的虚部符号来得到．

例 2　$-3 + 4i$ 的共轭复数是 $-3 - 4i$，记作 $\overline{-3 + 4i} = -3 - 4i$．■

注意如果 $z = a + bi$，那么

$$z\bar{z} = (a + bi)(a - bi) = a^2 - abi + bai - b^2i^2 = a^2 + b^2 \tag{3}$$

由于 $z\bar{z}$ 是实数且非负，它有一个平方根，z 的绝对值（或模）是实数 $|z|$ 且定义为

$$|z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}$$

如果 z 是实数，那么 $z = a + 0i$ 且 $|z| = \sqrt{a^2}$，它就等于平常意义下的 a 的绝对值．

下面列出一些关于共轭复数和绝对值的有用性质，设 w 和 z 表示复数.

1. $\bar{z}=z$ 的充分必要条件是 z 是实数.
2. $\overline{w+z}=\bar{w}+\bar{z}$.
3. $\overline{wz}=\bar{w}\,\bar{z}$ ；特别地，如果 r 是实数，则 $\overline{rz}=r\bar{z}$.
4. $z\bar{z}=|z|^2\geqslant 0$.
5. $|wz|=|w||z|$.
6. $|w+z|\leqslant|w|+|z|$.

如果 $z\neq 0$ ，那么 $|z|>0$ 且 z 有一个乘法的倒数，记作 $1/z$ 或 z^{-1} 且给出

$$\frac{1}{z}=z^{-1}=\frac{\bar{z}}{|z|^2}$$

当然，商 w/z 简单表示为 $w\cdot(1/z)$.

例 3 令 $w=3+4\mathrm{i}$ 和 $z=5-2\mathrm{i}$ ，计算 $z\bar{z},|z|$ 和 w/z .

解 由（3）式，

$$z\bar{z}=5^2+(-2)^2=25+4=29$$

对绝对值，$|z|=\sqrt{z\bar{z}}=\sqrt{29}$ ，为计算 w/z ，首先分子和分母同乘以 $\bar{z}$ ，即分母的共轭复数，由于（3）式的性质，这个运算可消去分母中的 i.

$$\frac{w}{z}=\frac{3+4\mathrm{i}}{5-2\mathrm{i}}=\frac{3+4\mathrm{i}}{5-2\mathrm{i}}\cdot\frac{5+2\mathrm{i}}{5+2\mathrm{i}}=\frac{15+6\mathrm{i}+20\mathrm{i}-8}{5^2+(-2)^2}$$

$$=\frac{7+26\mathrm{i}}{29}=\frac{7}{29}+\frac{26}{29}\mathrm{i}$$

■

几何解释

每一个复数 $z=a+b\mathrm{i}$ 对应平面 $\mathbb{R}^2$ 上的一个点 (a,b) ，如图 B-1 所示，水平轴称为**实轴**，因为它上面的点 $(a,0)$ 对应实数，垂直的轴称为**虚轴**，因为它上面的点 $(0,b)$ 对应. 形如 $0+b\mathrm{i}$ 的纯虚数，简记为 $b\mathrm{i}$，z 的共轭复数是 z 关于实轴的镜像，z 的绝对值是从原点到 (a,b) 的距离.

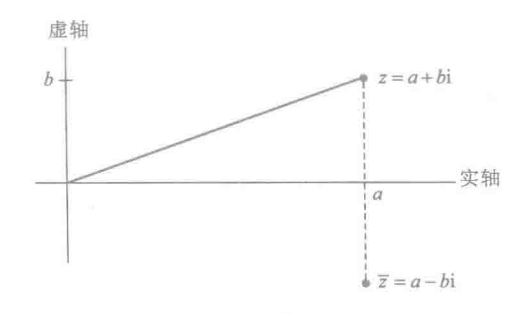


图 B-1 共轭复数是一个镜像

复数 $z=a+b\mathrm{i}$ 和 $w=c+b\mathrm{i}$ 的加法对应 $\mathbb{R}^2$ 中 (a,b) 和 (c,b) 的向量加法，如图 B-2 所示.

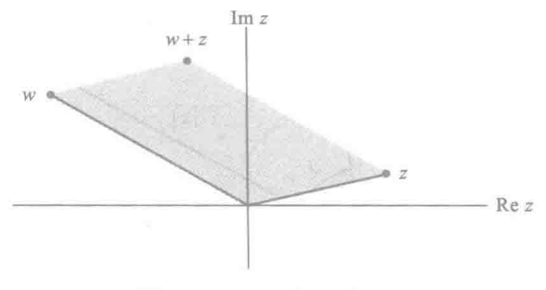


图 B-2　复数的加法

为给出复数乘法的图形表示，我们使用 $\mathbb{R}^2$ 中的极坐标，给定一个非零复数

$$z = a + b\mathrm{i}$$

设 φ 是正实轴和点 (a,b) 的夹角，如图 B-3 所示，此处 $-\pi < \varphi \leqslant \pi$，角 φ 称为 z 的**辐角**，我们记为 $\varphi = \arg z$，从三角函数得

$$a = |z|\cos\varphi \qquad b = |z|\sin\varphi$$

且

$$z = a + b\mathrm{i} = |z|(\cos\varphi + \mathrm{i}\sin\varphi)$$

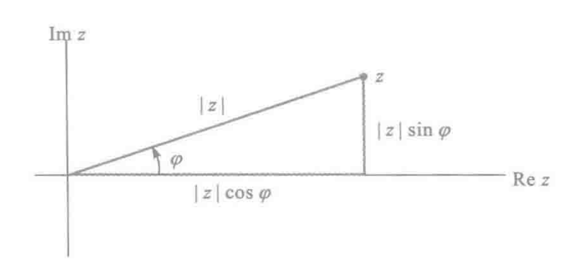


图 B-3　z 的极坐标

如果 w 是另一个非零复数，如

$$w = |w|(\cos\vartheta + \mathrm{i}\sin\vartheta)$$

利用关于正弦和余弦两角和的标准三角公式，可以验证

$$wz = |w||z|[\cos(\vartheta + \varphi) + \mathrm{i}\sin(\vartheta + \varphi)] \tag{4}$$

见图 B-4，类似可以写出极坐标形式的复数除法公式，乘法和除法公式可用文字描述如下.

> 极坐标系下，两个非零复数的乘法表示为它们绝对值的相乘和辐角相加. 极坐标系下，两个非零复数的除法表示为它们绝对值的相除和辐角之差.

例 4　a. 如果 w 的绝对值为 1，那么 $w = \cos\vartheta + \mathrm{i}\sin\vartheta$，此处 ϑ 是 w 的辐角，对任意非零复数 z 乘复数 w，简单表示为复数 z 旋转 ϑ 角度.

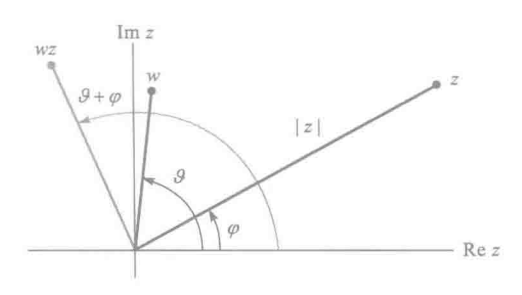


图 B-4 极坐标系中的乘法

b. i 本身的辐角为$\frac{\pi}{2}$，所以，用 i 来乘复数 z 是将 z 旋转$\frac{\pi}{2}$弧度，例如，$3+\mathrm{i}$旋转为$(3+\mathrm{i})\mathrm{i}=-1+3\mathrm{i}$．见图 B-5.

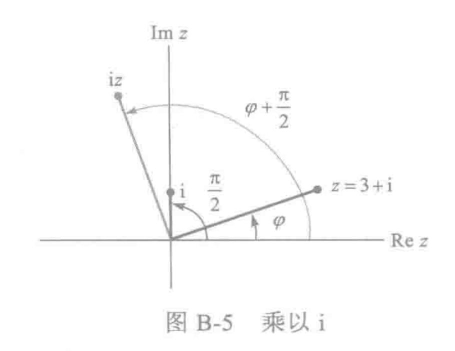


图 B-5 乘以 i

■

复数的幂

公式（4）中，如果 $z=w=r(\cos\varphi+\mathrm{i}\sin\varphi)$，则有

$$\begin{aligned}z^2&=r^2(\cos 2\varphi+\mathrm{i}\sin 2\varphi)\\z^3&=z\cdot z^2\\&=r(\cos\varphi+\mathrm{i}\sin\varphi)\cdot r^2(\cos 2\varphi+\mathrm{i}\sin 2\varphi)\\&=r^3(\cos 3\varphi+\mathrm{i}\sin 3\varphi)\end{aligned}$$

一般地，对任意正整数 k，

$$z^k=r^k(\cos k\varphi+\mathrm{i}\sin k\varphi)$$

这个结论被称为棣莫弗定理.

复数和 $\mathbb{R}^2$

尽管 $\mathbb{R}^2$ 中的元素和复平面 $\mathbb{C}$ 是一一对应的（见图 B-6 和图 B-7），并且加法实质上一致，但逻辑上 $\mathbb{R}^2$ 和 $\mathbb{C}$ 不同，在 $\mathbb{R}^2$ 上，我们仅能对一个向量作标量乘法，然而在 $\mathbb{C}$ 上，我们可以对任意两个复数相乘，并得到第三个复数.（$\mathbb{R}^2$ 中的点积不在考虑之内，原因是它产生一个数，此数不是 $\mathbb{R}^2$ 的元素.）我们用标量符号表示 $\mathbb{C}$ 中的元素，以强调这个区别.

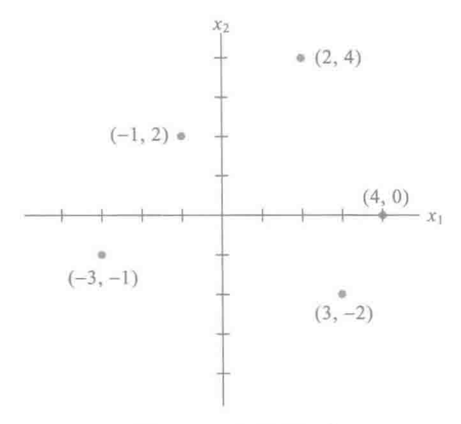


图 B-6　实平面 $\mathbb{R}^2$

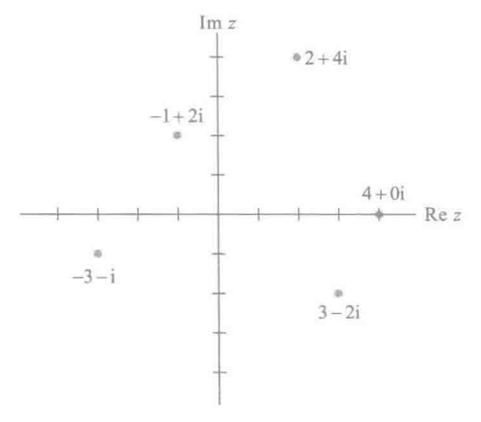


图 B-7　复平面 $\mathbb{C}$

术 语 表

A

adjugate (or classical adjoint) [**伴随矩阵**] 矩阵 adj A 是方阵 A 中 (i,j) 位置的元素，用 A 中 (i,j) 元素的代数余子式代替后，再通过转置得到的矩阵.

affine combination [**仿射组合**] 向量（$\mathbb{R}^n$ 中的点）的线性组合，且有关权重的和为 1.

affine dependence relation [**仿射相关关系**] 方程 $c_1\boldsymbol{v}_1+\cdots+c_p\boldsymbol{v}_p=\boldsymbol{0}$，这里权重 c_1，…，c_p 不全为 0，且 $c_1+\cdots+c_p=0$.

affine hull (或 affine span) of a set S [**集合 S 的仿射凸包（或仿射空间）**] S 中的点的所有仿射组合的集合，记为 aff S.

affine dependent set [**仿射相关集合**] $\mathbb{R}^n$ 中集合 $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p\}$ 存在实数 c_1，…，c_p 不全为 0，使得 $c_1+\cdots+c_p=0$ 且 $c_1\boldsymbol{v}_1+\cdots+c_p\boldsymbol{v}_p=\boldsymbol{0}$.

affine independent set [**仿射无关集合**] $\mathbb{R}^n$ 中集合 $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p\}$ 仿射不相关.

affine set (or affine subset) [**仿射集（或仿射子集）**] 若 $\boldsymbol{p},\boldsymbol{q}$ 是 S 中的点，那么对于每个实数 t，有 $(1-t)\boldsymbol{p}+t\boldsymbol{q}\in S$. 满足上述条件的集合 S 就是仿射集.

affine transformation [**仿射变换**] 一个形如 $T(\boldsymbol{x})=A\boldsymbol{x}+\boldsymbol{b}$ 的映射 $T:\mathbb{R}^n\to\mathbb{R}^m$，此处 A 是 $m\times n$ 阵且 $\boldsymbol{b}$ 属于 $\mathbb{R}^m$.

algebraic multiplicity [**代数重数**] 一个特征值的重数是作为特征方程的根的重数.

angle [**$\mathbb{R}^2$ 或 $\mathbb{R}^3$ 中非零向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 之间的角度**] 夹角 ϑ 是指从原点到点 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的两个线段之间的角度，通过点积联系起来.

$$\boldsymbol{u}\cdot\boldsymbol{v}=\|\boldsymbol{u}\|\|\boldsymbol{v}\|\cos\vartheta$$

associative law of multiplication [**乘法的结合律**] $A(BC)=(AB)C$，对所有 A,B,C 均成立.

attractor [**$\mathbb{R}^2$ 中一个动力系统的吸引子**] 当所有轨迹都趋于 0 时的原点.

augmented matrix [**增广矩阵**] 一个由线性方程组的系数矩阵和右边增加一列或多列的构成的矩阵，每一个增加的列包含给定系数矩阵对应方程组右边的常数项.

auxiliary equation [**辅助方程**] 一个关于变量 r 的多项式方程，来源于齐次差分方程的系数.

B

back-substitution (with matrix notation) [**双矩阵做回代变换**] 将阶梯形增广矩阵变换为简化形式的阶梯形矩阵时后一阶段的行变换，常用于求方程组的解.

backward phase (of row reduction) [**行变换的向后阶段**] 化简阶梯形矩阵为简化阶梯形矩阵算法的向后步骤.

band matrix [**带形矩阵**] 一个矩阵，它的非零元素位于主对角线的上下两侧的带形以内.

barycentric coordinates (of a point $\boldsymbol{p}$ with respect to an affinely independent set $S=\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\}$) [**（仿射无关集合 $S=\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\}$ 中点 $\boldsymbol{p}$ 的）重心坐标**] 唯一权重集 c_1，…，c_k，使得 $\boldsymbol{p}=c_1\boldsymbol{v}_1+\cdots+c_k\boldsymbol{v}_k$ 且 $c_1+\cdots+c_k=1$. （有时称为对于 S 的点 $\boldsymbol{p}$ 的仿射坐标.）

basic variable [基本变量] 线性方程组中对应系数矩阵主元列的变量.

basis (for a nontrivial subspace H of a vector space V) [向量空间 V 中的一个非平凡子空间 H 的基] 一个 V 中的向量集 $\mathcal{B}=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$，使得 (i) $\mathcal{B}$是线性无关集，(ii) 由$\mathcal{B}$生成的子空间和 H 一致，即 $H=\mathrm{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$.

$\mathcal{B}$-coordinates of $\boldsymbol{x}$ [$\boldsymbol{x}$ 的$\mathcal{B}$坐标] 见 $\boldsymbol{x}$ 相对于基$\mathcal{B}$的坐标.

best approximation [最佳逼近] 给定子空间中离给定向量的最近点.

bidiagonal matrix [两对角线矩阵] 一个非零元素位于主对角线和与主对角线相邻的一斜对角线上的矩阵.

block diagonal (matrix) [分块对角矩阵] 一个分块矩阵 $A=[A_{ij}]$，使得 $i\neq j$ 时，每一块 A_{ij} 是零矩阵.

block matrix [分块矩阵] 见矩阵分划.

block matrix multiplication [分块矩阵乘法] 分块后矩阵的行列乘法，将块作为数来对待.

block upper triangular (matrix) [上三角形分块矩阵] 分块矩阵 $A=[A_{ij}]$，使得一块 A_{ij} 对 $i>j$ 是零分块.

boundary point of a set S in $\mathbb{R}^n$ [$\mathbb{R}^n$ 中集合 S 的边界点] $\mathbb{R}^n$ 中以点 $\boldsymbol{p}$ 为心的开球与 S 的内外都有交点的点 $\boldsymbol{p}$.

bounded set in $\mathbb{R}^n$ [$\mathbb{R}^n$ 中的有界集] 对某个 $\delta>0$，包含在开球 $B(\boldsymbol{0},\delta)$ 中的一个集合.

$\mathcal{B}$-matrix (for T) [关于 T 的$\mathcal{B}$-矩阵] 对 V 中关于基$\mathcal{B}$的一个线性变换：$T:V\to V$ 对应的矩阵 $[T]_{\mathcal{B}}$，它具有如下的性质，对 V 中所有的 $\boldsymbol{x}$，$[T(\boldsymbol{x})]_{\mathcal{B}}=[T]_{\mathcal{B}}[\boldsymbol{x}]_{\mathcal{B}}$ 成立.

C

Cauchy-Schwarz inequality [柯西-施瓦茨不等式] 对所有 $\boldsymbol{u},\boldsymbol{v}$ 有 $|<\boldsymbol{u},\boldsymbol{v}>|\leqslant\|\boldsymbol{u}\|\cdot\|\boldsymbol{v}\|$.

change of basis [基的变换] 见矩阵坐标变换.

change-of-coordinates matrix (from a basis $\mathcal{B}$ to a basis $\mathcal{C}$) [从一个基$\mathcal{B}$到一个基$\mathcal{C}$ 的矩阵坐标变换] 一个矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 将$\mathcal{B}$坐标向量变换为$\mathcal{C}$坐标向量，$[\boldsymbol{x}]_{\mathcal{C}}=\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\boldsymbol{x}]_{\mathcal{B}}$，如果$\mathcal{C}$是 $\mathbb{R}^n$ 中的标准基，那么有时也将 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 记作 $P_{\mathcal{B}}$.

characteristic equation (of A) [A 的特征方程] $\det(A-\lambda I)=0$.

characteristic polynomial (of A) [A 的特征多项式] $\det(A-\lambda I)$ 或在有些教材中，$\det(\lambda I-A)$.

Cholesky factorization [楚列斯基分解] 一个分解 $A=R^{\mathrm{T}}R$，此处 R 是一个可逆上三角形矩阵，它的所有对角元素为正的.

closed ball (in $\mathbb{R}^n$) [($\mathbb{R}^n$ 中的) 闭球] $\mathbb{R}^n$ 中集合 $\{\boldsymbol{x}:\|\boldsymbol{x}-\boldsymbol{p}\|<\delta\}$，这里 $\boldsymbol{p}$ 在 $\mathbb{R}^n$ 中且 $\delta>0$.

closed set (in $\mathbb{R}^n$) [($\mathbb{R}^n$ 中的) 闭集] 包含所有的边界点的集合.

codomain (of a transformation $T:\mathbb{R}^n\to\mathbb{R}^m$) [变换 $T:\mathbb{R}^n\to\mathbb{R}^m$ 的余定义域] 集合 $\mathbb{R}^m$ 包含变换 T 的值域，通常，如果 T 映射一个向量空间 V 到另一个向量空间 W，那么 W 被称为 T 的余定义域.

coefficient matrix [系数矩阵] 一个矩阵，它的元素是一个线性方程组的系数.

cofactor [余子式] 一个数 $C_{ij}=(-1)^{i+j}\det A_{ij}$，称为矩阵 $A\,(i,j)$ 元素的余子式，此处 A_{ij} 是划去 A 中 i 行和 j 列后构成的子矩阵.

cofactor expansion [余子式展开] 一个计算 det A 的公式，利用 A 的一行或一列及其对应的余子式来展开，如按第一行展开：

$$\det A=a_{11}C_{11}+\cdots+a_{1n}C_{1n}$$

column-row expansion [列-行展开] AB 乘积的表达式作为外积之和，

$$\mathrm{col}_1(A)\mathrm{row}_1(B)+\cdots+\mathrm{col}_n(A)\mathrm{row}_n(B)$$

此处 n 是 A 的列数.

column space（of an $m\times n$ matrix A）[$m\times n$ 矩阵 A 的列空间] A 所有列的线性组合的集合 Col A，如果 $A=[\boldsymbol{a}_1\cdots\boldsymbol{a}_n]$，那么 $\text{Col } A=\text{Span}\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n\}$，等价地有 $\text{Col } A=\{\boldsymbol{y}:\boldsymbol{y}=A\boldsymbol{x}$, 对$\mathbb{R}^n$ 中的一些 $\boldsymbol{x}\}$.

column sum [列和] 一个矩阵中一列元素之和.

column vector [列向量] 只有一列的矩阵或具有几列的矩阵中的某一列.

commuting matrices [矩阵交换] 两个矩阵使得 $AB=BA$.

compact set （in $\mathbb{R}^n$）[（$\mathbb{R}^n$ 中的）紧集] $\mathbb{R}^n$ 中的集合既是闭的又是有界的.

companion matrix [友矩阵] 一类特殊形式的矩阵，其特征多项式是 $(-1)^n p(\lambda)$ 且 $p(\lambda)$ 是首项为 λ^n 的特殊多项式.

complex eigenvalue [复特征值] 一个 $n\times n$ 矩阵特征多项式的非实根.

complex eigenvector [复特征向量] $\mathbb{C}^n$ 中的非零向量 $\boldsymbol{x}$，使得 $A\boldsymbol{x}=\lambda\boldsymbol{x}$，此处 A 是一个 $n\times n$ 矩阵，λ是一个复特征值.

component of y orthogonal to $\boldsymbol{u}$（for $\boldsymbol{u}\neq\boldsymbol{0}$）[$\boldsymbol{y}$ 正交于 $\boldsymbol{u}(\boldsymbol{u}\neq\boldsymbol{0})$ 的分量] 向量 $\boldsymbol{y}-\dfrac{\boldsymbol{y}\cdot\boldsymbol{u}}{\boldsymbol{u}\cdot\boldsymbol{u}}\boldsymbol{u}$.

composition of linear transformations [线性变换的复合] 由连续应用两个或更多的线性变换产生的映射，如果这些变换是矩阵变换，例如，左乘 B 之后接着左乘 A，那么复合结果是映射 $\boldsymbol{x}\mapsto A(B\boldsymbol{x})$.

condition number（of A）[A 的条件数] 商 σ_1/σ_n，此处 σ_1 是 A 的最大特征值且 σ_n 是最小的特征值，如果 σ_n 是零，条件数是 $+\infty$.

conformable for block multiplication [与分块乘法相一致] 两个分块矩阵 A 和 B，其分块使得乘积 AB 有定义，即 A 的列分块等于 B 的行分块.

consistent linear system [相容的方程组] 至少具有一个解的一个线性方程组.

constrained optimization [条件优化] 当给出 $\boldsymbol{x}$ 的一个或更多限制时，一个量的最大化问题，例如满足条件 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 或 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{v}=0$，量 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 或 $\|A\boldsymbol{x}\|$ 的最大化问题.

consumption matrix [消费矩阵] 列昂惕夫投入产出模型中的一个矩阵，它的列是经济体系中各个部门的单位消耗向量.

contraction [压缩] 对一些数 r 做映射 $\boldsymbol{x}\mapsto r\boldsymbol{x}$，此处 $0\leqslant r\leqslant 1$.

controllable（pair of matrices）[可控制矩阵对] 一个矩阵对 (A,B)，A 是 $n\times n$ 矩阵，B 有 n 行并且满足

$$\text{rank}[B\ \ AB\ \ A^2B\cdots\ \ A^{n-1}B]=n$$

相关的是控制系统中的状态空间模型和差分方程 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k+B\boldsymbol{u}_k(k=0,1,\cdots)$.

convergent（sequence of vectors）[收敛向量序列] 一个序列 $\{\boldsymbol{x}_k\}$，如果 k 足够大，使得 $\boldsymbol{x}_k$ 中的元素可以接近某些固定向量中的元素.

convex combination（of points $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k$ in $\mathbb{R}^n$）[（$\mathbb{R}^n$ 中点 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k$ 的）凸组合] 向量（点）的线性组合，组合中的权重是非负的且权重和为 1.

convex hull（of a set S）[（集合 S 的）凸包] 在 S 中的点的所有凸组合集合. 记为 conv S.

convex set [凸集] 集合 S 具有如下特性：对于 S 的点 $\boldsymbol{p},\boldsymbol{q}$，线段 $\overline{\boldsymbol{pq}}$ 包含在 S 中.

coordinate mapping（determined by an ordered basis $\mathcal{B}$ in a vector space V）[一个向量空间的有序基所确定的坐标映射] 一个将 V 中每一个 $\boldsymbol{x}$ 与它的坐标向量 $[\boldsymbol{x}]_{\mathcal{B}}$ 联系起来的映射.

coordinates of $\boldsymbol{x}$ relative to the basis $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ [$\boldsymbol{x}$ 相对于基 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ 的坐标] 适合方程 $\boldsymbol{x}=c_1\boldsymbol{b}_1+\cdots+c_n\boldsymbol{b}_n$ 的系数 $c_1,\cdots,c_n$.

coordinate vector of $\boldsymbol{x}$ relative to $\mathcal{B}$ [$\boldsymbol{x}$ 相对于基$\mathcal{B}$的坐标向量] 向量 $[\boldsymbol{x}]_{\mathcal{B}}$，其中的元素是 $\boldsymbol{x}$ 在基$\mathcal{B}$下的坐标.

covariance (of variables x_i and x_j, for $i \neq j$) [变量 x_i 和变量 x_j 的协方差 $(i \neq j)$] 对一个观测矩阵的协方差矩阵 S 中的元素 s_{ij}，此处，x_i 和 x_j 分别遍历观测向量中所有 i 和 j 的坐标.

covariance matrix (or sample covariance matrix) [协方差矩阵（或样本协方差矩阵）] $p \times p$ 矩阵 S，定义为 $S = (N-1)^{-1}BB^{\mathrm{T}}$，此处 B 是一个 $p \times N$ 平均偏差形式的观测矩阵.

Cramer's Rule [克拉默法则] 当 A 是可逆矩阵时，公式给出方程组 $A\boldsymbol{x} = \boldsymbol{b}$ 的一个解 $\boldsymbol{x}$ 中的每个元素取值.

cross-product term [交叉乘积项] 二次型中的项 cx_ix_j，其中 $i \neq j$.

cube [立体] 由 6 个平面界定的三维固体，且三个面共用一个顶点.

D

decoupled system [解耦系统] 一个差分方程 $\boldsymbol{y}_{k+1} = A\boldsymbol{y}_k$，或微分方程 $\boldsymbol{y}'(t) = A\boldsymbol{y}(t)$，此处 A 是对角矩阵，离散变化方程中 $\boldsymbol{y}_k$（作为一个关于 k 的函数）的每个分量或连续变化方程中向量值函数 $\boldsymbol{y}(t)$ 的每个分量，当 $k \to \infty$ 或 $t \to \infty$ 时，不受其他分量变化的影响.

design matrix [设计矩阵] 在线性模型 $\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ 中的矩阵 X，此处 X 的列，以某种方式由一些独立变量的观测值所确定.

determinant (of a square matrix A) [一个方阵 A 的行列式] 数 det A 归纳定义为 A 的第一行余子式的展开. 也等于 $(-1)^r$ 乘以由 A 作行变换而得到的任何阶梯矩阵 U 中对角元素的乘积，该阶梯矩阵是通过行替换和 r 次行交换得到（但没有行的倍乘变换）.

diagonal entries (in a matrix) [一个矩阵的对角元素] 元素具有相同的行下标和列下标.

diagonalizable (matrix) [可对角化矩阵] 一个矩阵可以写成分解形式 PDP^{-1}，此处 D 是一个对角矩阵，P 是一个可逆矩阵.

diagonal matrix [对角矩阵] 一个方阵，不在主对角线上的所有元素为零.

difference equation (or linear recurrence relation) [差分方程（或线性回归方程）] 一个形如 $\boldsymbol{x}_{k+1} = A\boldsymbol{x}_k (k = 0,1,2,\cdots)$ 的方程，它的解是一个向量序列，$\boldsymbol{x}_0, \boldsymbol{x}_1, \cdots$.

dilation [拉伸变换] 对标量 r，作变换 $\boldsymbol{x} \mapsto r\boldsymbol{x}$，其中 $r > 1$.

dimension (of a vector space V) [向量空间 V 的维数] V 的一个基中向量的个数，记作 dim V，零空间的维数为零.

discrete linear dynamical system [离散线性动力系统] 一个形如 $\boldsymbol{x}_{k+1} = A\boldsymbol{x}_k$ 的差分方程，描述随时间演变的系统变化（常指物理系统），物理系统用离散时刻来度量，当 $k = 0,1,2,\cdots$，且系统在时刻 k 时的状态是一个向量，其元素给出所感兴趣的事实.

distance between $\boldsymbol{u}$ and $\boldsymbol{v}$ [$\boldsymbol{u}$ 和 $\boldsymbol{v}$ 之间的距离] 向量 $\boldsymbol{u} - \boldsymbol{v}$ 的长度，记作 $\mathrm{dist}(\boldsymbol{u}, \boldsymbol{v})$.

distance to a subspace [到一个子空间的距离] 从一个给定点（向量）$\boldsymbol{v}$ 到一个子空间的最近点的距离.

distributive laws [分配律] 对所有 A, B, C，（左）$A(B + C) = AB + AC$ 和（右）$(B + C)A = BA + CA$

domain (of a transformation T) [变换 T 的定义域] $T(\boldsymbol{x})$ 有定义的向量 $\boldsymbol{x}$ 的集合.

dot product [点积] 见内积.

dynamical system [动力系统] 见离散线性动力系统.

E

echelon form (or row echelon form, of a matrix) [矩阵的阶梯形式（或行阶梯形式）] 一个阶

梯矩阵行等价于给定矩阵.

echelon matrix（or row echelon matrix）[**阶梯矩阵（或行阶梯矩阵）**] 一个长方形的矩阵具有三个特性：(1) 所有非零行位于所有零行的上方，(2) 每一个一行中的主元素所在列处于上行主元素所在列的右边. (3) 在主元素之下的所有同一列的元素是零.

eigenfunctions（of a differential equation $x'(t)=Ax(t)$）[**一个微分方程 $x'(t)=Ax(t)$ 的特征函数**] 一个函数 $x(t)=ve^{\lambda t}$，此处 v 是 A 的特征向量且 λ 是对应的特征值.

eigenspace（of A corresponding to λ）[**矩阵 A 对应 λ 的特征子空间**] $Ax=\lambda x$ 所有解的集合，它包含零向量和所有对应 λ 的特征向量，此处 λ 是 A 的一个特征值.

eigenvalue（of A）[**A 的特征值**] 一个数 λ，使得方程 $Ax=\lambda x$ 有一个非零向量解 x.

eigenvector（of A）[**A 的特征向量**] 一个非零向量 x，使得 $Ax=\lambda x$ 对一些数 λ 成立.

eigenvector basis [**特征向量基**] 包含给定矩阵的全部特征向量的基.

eigenvector decomposition（of x）[**x 的特征向量分解**] 一个方程，$x=c_1v_1+\cdots+c_nv_n$，表示 x 作为矩阵特征向量的线性组合.

elementary matrix [**初等矩阵**] 一个可逆矩阵，它是对单位矩阵作一次基本行变换得到的矩阵.

elementary row operations [**初等行变换**] (1)（替换）用自己所在行与其他行的倍数之和替换一行. (2) 交换两行. (3)（倍乘）用非零数乘某行的所有元素.

equal vectors [**相等向量**] $\mathbb{R}^n$ 中对应的元素相同的向量.

equilibrium prices [**平衡价格**] 经济体系中各个部门总支出的价格集合，使得每个部门的总收入与总支出正好均衡.

equilibrium vector [**平衡向量**] 见稳态向量.

equivalent（linear）systems [**等价的线性系统**] 具有同样解集的线性系统.

exchange model [**交换模型**] 见列昂惕夫交换模型.

existence question [**存在性问题**] “一个系统的解是否存在？”，即“系统是否相容？”，也是“对所有 b，$Ax=b$ 的一个解是否存在？”

expansion by cofactors [**用余子式展开**] 见余子式展开.

explicit description（of a subspace W of $\mathbb{R}^n$）[**显式描述 $\mathbb{R}^n$ 中一个子空间 W**] 用参数表示 W，它是一个特殊向量集合的所有线性组合的全体.

extreme point（of a convex set S）[**（凸集 S 的）端点**] 存在 S 中的点 p，使得 p 不在位于 S 的任何线段内部.（即若 x 与 y 属于 S，且 p 在线段 $\overline{xy}$ 上，则 $p=x$ 或 $p=y$.）

F

factorization（of A）[**A 的分解**] 将 A 表示为两个或更多矩阵乘积的一个方程.

final demand vector（or bill of final demands）[**最终需求向量（或最终需求的清单）**] Leontief 投入-产出模型中的向量 d，列出部分经济体系中非生产性的各部门货物和服务的价值. 向量 d 可以表示消费者需求、政府消费、生产者剩余、出口或外部需求.

finite-dimensional（vector space）[**有限维向量空间**] 一个由有限个向量的集合所生成的向量空间.

flexibility matrix [**弹性矩阵**] 一个矩阵，当单位大小的力作用在梁的第 j 个点时，它的第 j 列给出该弹性梁在特定点处的弯曲.

floating point arithmetic [**浮点算术运算**] 数字用十进制 $\pm 0.d_1\cdots d_p\times 10^r$ 表示，此处 r 是一个整数且表示小数点右边的位字，数 p 常常位于 8 和 16 之间.

flop [**浮运算**] 两个实浮点数的一次算术运算(+, −, *, /).

forward phase (of row reduction) [**行简化的向前阶段**] 简化一个矩阵为阶梯矩阵的第一部分算法.

Fourier approximation (of order n) [**n 阶傅里叶逼近**] 在 n 阶三角多项式子空间中，与空间 $C[0,2\pi]$ 中给定函数的距离最近的点.

Fourier coefficients [**傅里叶系数**] 在傅里叶逼近一个函数中，三角多项式的系数.

Fourier series [**傅里叶级数**] 一个无穷级数，在内积空间 $C[0,2\pi]$ 内收敛于一个函数，其内积由一个定积分确定.

free variable [**自由变量**] 一个线性方程组中不是基本变量的任意变量.

full rank (matrix) [**满秩矩阵**] 一个 $m\times n$ 矩阵，它的秩等于 m 和 n 的最小值.

fundamental set of solutions [**基础解系**] 齐次线性方程或齐次微分方程所有解构成的集合的一个基.

fundamental subspaces (determined by A) [**A 的基础子空间**] A 的零空间和 A 的列空间以及 A^{T} 的零空间和 A^{T} 的列空间，Col A^{T} 常称为 A 的行空间.

G

Gaussian elimination [**高斯消元**] 见行化简方法.

general least-squares problem [**一般的最小二乘问题**] 给定一个 $m\times n$ 矩阵 A 和一个属于 $\mathbb{R}^m$ 的向量 $\boldsymbol{b}$，求属于 $\mathbb{R}^n$ 的 $\hat{\boldsymbol{x}}$，使得 $\|\boldsymbol{b}-A\hat{\boldsymbol{x}}\|\leqslant\|\boldsymbol{b}-A\boldsymbol{x}\|$ 对所有 $\mathbb{R}^n$ 中的 $\boldsymbol{x}$ 成立.

general solution (of a linear system) [**一个线性方程组的通解**] 参数表示的解集合，它用任意参数形式的自由变量表示基本变量，在 1.5 节之后，参数表示被写成向量形式.

Givens rotation [**吉温斯旋转**] 计算机中使用的从 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 的线性变换，其作用是产生一个向量（常指矩阵的一列）中的零元素.

Gram matrix (of A) [**矩阵 A 的格拉姆矩阵**] 矩阵 $A^{\mathrm{T}}A$.

Gram-Schmidt process [**格拉姆-施密特方法**] 对生成子空间的给定向量集合，生成正交或单位正交基的一个算法.

H

homogeneous coordinates [**齐次坐标**] 在 $\mathbb{R}^3$ 中，对任何 $H\neq 0$，将 (x,y,z) 表示为 (X,Y,Z,H)，此处 $x=X/H, y=Y/H$ 和 $z=Z/H$. 在 $\mathbb{R}^2$ 中，H 常取作 1，(x,y) 的齐次坐标常写成 $(x,y,1)$.

homogeneous equation [**齐次方程**] 一个形如 $A\boldsymbol{x}=\boldsymbol{0}$ 的方程，可能写成是一个向量方程或一个线性方程组.

homogeneous form of (a vector) $\boldsymbol{v}$ in $\mathbb{R}^n$ [**$\mathbb{R}^n$ 中（向量）$\boldsymbol{v}$ 的齐次形式**] $\mathbb{R}^{n+1}$ 中的点 $\tilde{\boldsymbol{v}}=\begin{bmatrix}\boldsymbol{v}\\1\end{bmatrix}$.

Householder reflection [**豪斯霍尔德镜像**] 一个变换 $\boldsymbol{x}\mapsto Q\boldsymbol{x}$，此处 $Q=I-2\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}$ 且 $\boldsymbol{u}$ 是一个单位向量 $(\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}=1)$.

hyperplane (in $\mathbb{R}^n$) [**（$\mathbb{R}^n$ 中的）超平面**] $\mathbb{R}^n$ 中维数为 $n-1$ 的一个平面. 也是一个 $n-1$ 维子空间的平移.

I

identity matrix (denoted by I or I_n) [**单位矩阵（记作 I 或 I_n）**] 一个对角线上为 1、其他元素为 0 的方阵.

ill-conditioned matrix [**病态矩阵**] 一个具有大的（或者无穷大）条件数的矩阵；如果矩阵中的一些元素改变一点，矩阵是奇异矩阵或会变成

奇异矩阵.

image（of a vector $\boldsymbol{x}$ under a transformation T）[一个向量在一个变换 T 下的像] 用 T 指定给 $\boldsymbol{x}$ 的向量 $T(\boldsymbol{x})$.

implicit description（of a subspace W of $\mathbb{R}^n$）[$\mathbb{R}^n$ 的一个子空间 W 的隐式描述] 一个或多个齐次方程的解集合描述 W 中点的特性.

Im $\boldsymbol{x}$ [$\boldsymbol{x}$ 的虚部] 由 $\mathbb{C}^n$ 中向量 $\boldsymbol{x}$ 的虚部元素所形成的 $\mathbb{R}^n$ 中的向量.

inconsistent linear system [不相容线性方程组] 一个没有解的线性方程组.

indefinite matrix [不定矩阵] 一个对称矩阵 A，使得 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的值既有正值又有负值.

indefinite quadratic form [不定二次型] 一个二次型 Q，使得 $Q(\boldsymbol{x})$既有正值又有负值.

infinite-dimensional（vector space）[无限维向量空间] 一个非零向量空间 V 没有有限基.

inner product [内积] 数量 $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}$，常写成 $\boldsymbol{u}\cdot\boldsymbol{v}$，此处 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是 $\mathbb{R}^n$ 中的向量，作为 $n\times1$ 矩阵，也称为 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的点积. 一般地，一个向量空间中的函数，给出每一对向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 对应的一个数 $\langle\boldsymbol{u},\boldsymbol{v}\rangle$，它满足一些公理，见 6.7 节.

inner product space [内积空间] 一个已定义了内积的向量空间.

input-output matrix [投入产出矩阵] 见消费矩阵.

input-output model [投入产出模型] 见列昂惕夫投入产出模型.

interior point（of a set S in $\mathbb{R}^n$）[（$\mathbb{R}^n$ 中集合 S 的）内点] S 中的一个点 $\boldsymbol{p}$，使得对某个 $\delta>0$，中心为 $\boldsymbol{p}$ 的开球 $B(\boldsymbol{p},\delta)$ 包含于 S 中.

intermediate demands [中间需求] 对货物或服务的需求会被消耗在其他货物生产过程中或顾客的服务中，如果 $\boldsymbol{x}$ 是生产水平且 C 是消费矩阵，那么 $C\boldsymbol{x}$ 列出中间需求.

interpolating polynomial [插值多项式] 一个多项式，其图像通过 $\mathbb{R}^2$ 中的一个点集的每一个点.

invariant subspace（for A）[A 的不变子空间] 一个子空间 H，使得当 $\boldsymbol{x}\in H$ 时，$A\boldsymbol{x}$ 仍然属于 H.

inverse（of an $n\times n$ matrix A）[一个 $n\times n$ 矩阵 A 的逆] 一个 $n\times n$ 矩阵 A^{-1} 使得 $AA^{-1}=A^{-1}A=I_n$.

inverse power method [逆幂法] 当具有λ的一个好的初始估计时，一个估计方阵特征值λ的算法.

invertible linear transformation [可逆线性变换] 一个线性变换 $T:\mathbb{R}^n\to\mathbb{R}^n$，使得存在一个函数 $S:\mathbb{R}^n\to\mathbb{R}^n$ 对所有 $\boldsymbol{x}\in\mathbb{R}^n$，满足 $T(S(\boldsymbol{x}))=\boldsymbol{x}$ 和 $S(T(\boldsymbol{x}))=\boldsymbol{x}$.

invertible matrix [可逆的矩阵] 一个具有逆阵的方阵.

isomorphic vector spaces [同构向量空间] 两个向量空间 V 和 W，且存在一个一对一线性变换 T，将 V 映上到 W.

isomorphism [同构] 从一个向量空间到另一个向量空间的一个一对一线性映射.

K

kernel（of a linear transformation $T:V\to W$）[一个线性变换 $T:V\to W$ 的核] V 中满足 $T(\boldsymbol{x})=\boldsymbol{0}$ 的所有 $\boldsymbol{x}$ 的集合.

Kirchhoff's Laws [基尔霍夫定律] （1）（电压定律）环路中一个方向上的电压降 RI 的代数和等于环路中同一方向电源电压的代数之和.（2）（电流定律）一个分支中的电流是流向该分支环路电流的代数和.

L

ladder network [梯形网络] 一个由两个或更多电路串联而成的电路网络.

leading entry [首项元素] 一个矩阵一行中最左

边的非零元素.

least-squares error [最小二乘误差] 从 $\boldsymbol{b}$ 到 $A\hat{\boldsymbol{x}}$ 的距离 $\|\boldsymbol{b}-A\hat{\boldsymbol{x}}\|$，其中 $\hat{\boldsymbol{x}}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的一个最小二乘解.

least-squares line [最小二乘直线] 直线 $y=\hat{\beta}_0+\hat{\beta}_1 x$，使得方程 $\boldsymbol{y}=X\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ 的最小二乘误差达到最小.

least-squares solution (of $A\boldsymbol{x}=\boldsymbol{b}$) [方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解] 一个向量 $\hat{\boldsymbol{x}}$ 使得对所有属于 $\mathbb{R}^n$ 的 $\boldsymbol{x}$ 有

$$\|\boldsymbol{b}-A\hat{\boldsymbol{x}}\|\leqslant\|\boldsymbol{b}-A\boldsymbol{x}\|$$

left inverse (of A) [A 的左逆] 任何矩形矩阵 C 使得 $CA=I$.

left-multiplication (by A) [用 A 左乘] 用 A 左乘一个向量或一个矩阵.

left singular vectors (of A) [A 的左奇异向量] 在奇异值分解 $A=U\Sigma V^{\mathrm{T}}$ 中的 U 的列.

length (or norm of $\boldsymbol{v}$) [$\boldsymbol{v}$ 的长度或范数] 数 $\|\boldsymbol{v}\|=\sqrt{\boldsymbol{v}\cdot\boldsymbol{v}}=\sqrt{\langle\boldsymbol{v},\boldsymbol{v}\rangle}$.

Leontief exchange (or closed) model [列昂惕夫交换（或封闭）模型] 经济体系中的一个投入和产出都固定的模型，且部门产出的一组价格，要求满足每个部门的收入等于该部门的支出. 这个"平衡"条件表示为一个价格未知的线性方程组.

Leontief input-output model (or Leontief production equation) [列昂惕夫投入产出模型（或列昂惕夫产量方程）] 方程 $\boldsymbol{x}=C\boldsymbol{x}+\boldsymbol{d}$，其中 $\boldsymbol{x}$ 是产量，$\boldsymbol{d}$ 是最终需求，C 是消费（或投入产出）矩阵. C 的第 j 列给出该部门第 j 个顾客每单位产出的投入.

level set (or gradient) of a linear functional f on $\mathbb{R}^n$ [$\mathbb{R}^n$ 上线性函数 f 的水平集（或梯度）] 集合 $[f:d]=\left\{\boldsymbol{x}\in\mathbb{R}^n : f(\boldsymbol{x})=d\right\}$.

linear combination [线性组合] 向量标量乘法之和，其中标量称为权值.

linear dependence relation [线性相关关系] 一个齐次向量方程具有特殊的权值系数，且至少一个权值不为零.

linear equation (in the variables $x_1,\cdots,x_n$) [变量 $x_1,\cdots,x_n$ 的线性方程] 一个可以写成形如 $a_1x_1+a_2x_2+\cdots+a_nx_n=b$ 的方程，此处 b 和系数 $a_1,\cdots,a_n$ 是实或复数.

linear filter [线性滤波] 一个线性差分方程用于变换离散时间信号.

linear functional (on $\mathbb{R}^n$) [（$\mathbb{R}^n$ 上的）线性函数] 从 $\mathbb{R}^n$ 到 $\mathbb{R}$ 的线性变换 f.

linearly dependent (vectors) [线性相关向量] 带标号向量的集合 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 具有性质，存在不全为零的权值 $c_1,\cdots,c_p$，使得 $c_1\boldsymbol{v}_1+\cdots+c_p\boldsymbol{v}_p=\boldsymbol{0}$. 即向量方程 $c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_p\boldsymbol{v}_p=\boldsymbol{0}$ 有一组非零解.

linearly independent (vectors) [线性无关向量] 带标号向量的集合 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 具有性质，向量方程 $c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_p\boldsymbol{v}_p=\boldsymbol{0}$ 只有平凡解 $c_1=\cdots=c_p=0$.

linear model (in statistics) [统计中的线性模型] 任何形如 $\boldsymbol{y}=X\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ 的方程，此处 X 和 $\boldsymbol{y}$ 是已知的，$\boldsymbol{\beta}$ 的选取，使得剩余向量 $\boldsymbol{\varepsilon}$ 的长度最小.

linear system [线性系统] 一个或多个包含同样变量如 $x_1,\cdots,x_n$ 的线性方程集合.

linear transformation T (from a vector space V into a vector space W) [从向量空间 V 到向量空间 W 的线性变换 T] 一个法则 T 指定 V 中任一向量 $\boldsymbol{x}$ 对应 W 中唯一向量 $T(\boldsymbol{x})$，使得（i）对任意 V 中的向量 $\boldsymbol{u},\boldsymbol{v}$，有 $T(\boldsymbol{u}+\boldsymbol{v})=T(\boldsymbol{u})+T(\boldsymbol{v})$. 且（ii）对 V 中所有向量 $\boldsymbol{u}$ 和所有数 c，有 $T(c\boldsymbol{u})=cT(\boldsymbol{u})$. 记号：$T:V\to W$；当 $T:\mathbb{R}^n\to\mathbb{R}^m$ 和 A 是 T 的标准矩阵时，也可表示为 $\boldsymbol{x}\mapsto A\boldsymbol{x}$.

line through p parallel to v [通过 p 且平行 v 的直线] 集合 $\{p+tv: t\in\mathbb{R}\}$.

loop current [闭路电流] 流过一个回路电流的总和，使得回路中 RI 电压降的代数和等于回路中电源电压的代数和.

lower triangular matrix [下三角形矩阵] 一个对角线以上的元素全为零的矩阵.

lower triangular part (of A) [A 的下三角形部分] 一个下三角形矩阵，它的位于主对角线和对角线以下的元素与 A 中同样位置的元素相同.

LU factorization [LU 分解] 一个矩阵 A 形如 $A=LU$ 的表示，此处 L 是对角线上元素为 1 的下三角形方阵（一个单位下三角形矩阵），U 是 A 的一个阶梯矩阵.

M

magnitude (of a vector) [一个向量的长度] 见范数.

main diagonal (of a matrix) [一个矩阵的主对角线] 具有相同行下标和列下标的元素.

mapping [映射] 见变换.

Markov chain [马尔可夫链] 一个概率向量序列 $\boldsymbol{x}_0,\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots$ 连同随机矩阵 P，使得 $\boldsymbol{x}_{k+1}=P\boldsymbol{x}_k$ 对任意 $k=0,1,2,\cdots$，成立.

matrix [矩阵] 一个矩形数组.

matrix equation [矩阵方程] 一个至少包含一个矩阵的方程，例如，$A\boldsymbol{x}=\boldsymbol{b}$.

matrix for T relative to bases $\mathcal{B}$ and $\mathcal{C}$ [相对于基 $\mathcal{B}$ 和基 $\mathcal{C}$ 的矩阵 T] 对一个线性变换 $T:V\to W$ 的矩阵 M，具有性质，对任何属于 V 的 $\boldsymbol{x}$，有 $[T(\boldsymbol{x})]_{\mathcal{C}}=M[\boldsymbol{x}]_{\mathcal{B}}$，此处 $\mathcal{B}$ 是 V 的基，$\mathcal{C}$ 是 W 的基. 当 $W=V$ 和 $\mathcal{C}=\mathcal{B}$ 时，矩阵 M 被称为 T 的 $\mathcal{B}$ 矩阵，且记为 $[T]_{\mathcal{B}}$.

matrix of observations [观测矩阵] 一个 $p\times N$ 矩阵，它的列是观测向量，每一列给出一个特定总体或集合中的个体或对象的 p 个测量值.

matrix transformation [矩阵变换] 一个映射 $\boldsymbol{x}\mapsto A\boldsymbol{x}$，此处 A 是一个 $m\times n$ 矩阵，$\boldsymbol{x}$ 表示 $\mathbb{R}^n$ 中的任意向量.

maximal linearly independent set (in V) [V 中最大的线性无关集] V 中一个线性无关集合 $\mathcal{B}$，使得属于 V，但不属于 $\mathcal{B}$ 的一个向量 V 添加到 $\mathcal{B}$ 中，则新集合是线性相关的.

mean-deviation form (of a matrix of observations) [一个观测矩阵的平均偏差形式] 一个矩阵，它的行向量是平均偏差形式，任一行元素之和为零.

mean-deviation form (of a vector) [一个向量的平均偏差形式] 一个元素之和为零的向量.

mean square error [均方误差] 内积空间中一个逼近的误差，此处的内积由定积分来定义.

migration matrix [迁移矩阵] 一个矩阵给出不同位置间从一个阶段到下一阶段移动的百分比.

minimal spanning set (for a subspace H) [一个子空间 H 的最小生成集] 集合 $\mathcal{B}$ 生成 H 并且具有性质，如果 $\mathcal{B}$ 中的任一元素从 $\mathcal{B}$ 中去掉，则新的集合不能生成 H.

$m\times n$ matrix [$m\times n$ 矩阵] 一个具有 m 行和 n 列的矩阵.

Moore-Penrose inverse [穆尔-彭罗斯逆] 见伪逆.

multiple regression [多元回归] 一个线性模型包含多个线性无关变量和一个相关变量.

N

nearly singular matrix [几乎奇异矩阵] 一个病态矩阵.

negative definite matrix [负定矩阵] 一个对称矩阵 A，使得对所有 $\boldsymbol{x}\neq\boldsymbol{0}$，有 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}<0$ 成立.

negative definite quadratic form [负定二次型] 一个二次型 Q，使得对所有 $\boldsymbol{x}\neq\boldsymbol{0}$，有 $Q(\boldsymbol{x})<0$

成立.

negative semidefinite matrix [半负定矩阵] 一个对称矩阵 A，使得对所有 $\boldsymbol{x}$，有 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} \leqslant 0$ 成立.

negative semidefinite quadratic form [半负定二次型] 一个二次型 Q，使得对所有 $\boldsymbol{x}$，$Q(\boldsymbol{x}) \leqslant 0$ 成立.

nonhomogeneous equation [非齐次方程] 一个形如 $A\boldsymbol{x}=\boldsymbol{b}$ 且 $\boldsymbol{b} \neq \boldsymbol{0}$ 的方程，也可以表示一个向量方程或一个线性方程组.

nonsingular (matrix) [非奇异矩阵] 一个可逆矩阵.

nontrivial solution [非平凡解] 齐次方程或齐次线性方程组的一个非零解.

nonzero (matrix or vector) [非零（矩阵或向量）] 一个矩阵（可能仅有一行或仅有一列）包含至少一个非零元素.

norm (or length of $\boldsymbol{v}$) [$\boldsymbol{v}$ 的范数或长度] 标量 $\|\boldsymbol{v}\|=\sqrt{\boldsymbol{v}\cdot\boldsymbol{v}}=\sqrt{\langle \boldsymbol{v}\cdot\boldsymbol{v}\rangle}$.

normal equations [标准方程] 表示为 $A^{\mathrm{T}}A\boldsymbol{x}=A^{\mathrm{T}}\boldsymbol{b}$ 的方程，它的解给出 $A\boldsymbol{x}=\boldsymbol{b}$ 的所有最小二乘解.

在统计学中，常见的记号是 $X^{\mathrm{T}}X\boldsymbol{\beta}=X^{\mathrm{T}}\boldsymbol{y}$.

normalizing (a nonzero vector $\boldsymbol{v}$) [一个非零向量的单位化] 用 $\boldsymbol{v}$ 的正倍数得到一个单位向量 $\boldsymbol{u}$ 的过程.

null space (of an $m\times n$ matrix A) [一个 $m\times n$ 矩阵 A 的零空间] 齐次方程 $A\boldsymbol{x}=\boldsymbol{0}$ 的所有解的集合 Nul A. Nul $A=\{\boldsymbol{x}:\boldsymbol{x}$ 属于$\mathbb{R}^n$且$A\boldsymbol{x}=\boldsymbol{0}\}$.

O

observation vector [观测向量] 线性模型 $\boldsymbol{y}=X\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ 中的向量 $\boldsymbol{y}$，此处 $\boldsymbol{y}$ 的元素是因变量的观测值.

one-to-one (mapping) [一对一映射] 一个映射 $T:\mathbb{R}^n\to\mathbb{R}^m$，使得 $\mathbb{R}^m$ 中的每一个 b 是 $\mathbb{R}^n$ 中最多一个元素 $\boldsymbol{x}$ 的像.

onto (mapping) [映上映射] 一个映射 $T:\mathbb{R}^n\to\mathbb{R}^m$，使得 $\mathbb{R}^m$ 中的每一个元素是 $\mathbb{R}^n$ 中至少一个元素 $\boldsymbol{x}$ 的像.

open ball $\mathrm{B}(\boldsymbol{p},\delta)$ in $\mathbb{R}^n$ [$\mathbb{R}^n$ 中的开球 $\mathrm{B}(\boldsymbol{p},\delta)$] $\mathbb{R}^n$ 中的集合 $\{\boldsymbol{x}:\|\boldsymbol{x}-\boldsymbol{p}\|<\delta\}$，这里 $\delta>0$.

open set S in $\mathbb{R}^n$ [$\mathbb{R}^n$ 中的开集 S] 不包含边界点的集合.（等价于，若 S 中的每个点都是内点，则 S 是开的.）

origin [原点] 零向量.

orthogonal basis [正交基] 一个基，它也是一个正交集合.

orthogonal complement (of W) [W 的正交补] 与 W 中所有向量都正交的集合 $W^{\perp}$.

orthogonal decomposition [正交分解] 一个向量 $\boldsymbol{y}$ 表示为两向量之和，一个向量在特定的子空间 W 中，另一向量在空间 $W^{\perp}$. 一般地，分解 $\boldsymbol{y}=c_1\boldsymbol{u}_1+\cdots+c_p\boldsymbol{u}_p$，这里 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是包含 $\boldsymbol{y}$ 的子空间的一个正交基.

orthogonally diagonalizable (matrix) [可正交对角化矩阵] 矩阵 A 有一个分解 $A=PDP^{-1}$，其中 P 是正交矩阵 $(P^{-1}=P^{\mathrm{T}})$ 且 D 是对角矩阵.

orthogonal matrix [正交矩阵] 一个可逆方阵 U，使得 $U^{-1}=U^{\mathrm{T}}$.

orthogonal projection of $\boldsymbol{y}$ onto $\boldsymbol{u}$ (or onto the line through $\boldsymbol{u}$ and the origin, for $\boldsymbol{u}\neq\boldsymbol{0}$) [$\boldsymbol{y}$ 在 $\boldsymbol{u}$ 上的正交投影（或在通过 $\boldsymbol{u}$ 和原点的直线上的投影，其中 $\boldsymbol{u}\neq\boldsymbol{0}$）] 向量 $\hat{y}$ 定义为 $\hat{\boldsymbol{y}}=\dfrac{\boldsymbol{y}\cdot\boldsymbol{u}}{\boldsymbol{u}\cdot\boldsymbol{u}}\boldsymbol{u}$.

orthogonal projection of $\boldsymbol{y}$ onto W [$\boldsymbol{y}$ 在 W 上的正交投影] W 中的唯一向量 $\hat{\boldsymbol{y}}$，使得 $\boldsymbol{y}-\hat{\boldsymbol{y}}$ 与 W 正交，记号 $\hat{\boldsymbol{y}}=\mathrm{proj}_W\boldsymbol{y}$.

orthogonal set [正交集] 向量 S 的集合，使得对 S 中的不同向量 $\boldsymbol{u},\boldsymbol{v}$，有 $\boldsymbol{u}\cdot\boldsymbol{v}=0$ 成立.

orthogonal to W [与 W 正交] 与 W 中任一向量正交.

orthonormal basis [单位正交基] 一个由单位正交向量集合构成的基.

orthonormal set [单位正交集] 一个由单位向量构成的正交集合.

outer product [外积] 一个矩阵乘积 $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$，此处 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是 $\mathbb{R}^n$ 中作为 $n\times 1$ 矩阵的向量.（转置符号在符号 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的“外侧”）.

overdetermined system [超定方程组] 一个线性方程组中，方程的个数多于未知变量的个数.

P

parallel flats [平行平面] 两个或更多的平面，满足每个平面是其他平面的平移.

parallelogram rule for addition [加法的平行四边形法则] 两个向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 之和是 $\boldsymbol{u},\boldsymbol{v}$ 和 $\mathbf{0}$ 确定的平行四边形的对角线的几何解释.

parameter vector [参数向量] 线性模型 $\boldsymbol{y}=X\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ 中的未知向量 $\boldsymbol{\beta}$.

parametric equation of a line [一条直线的参数方程] 一个形如 $\boldsymbol{x}=\boldsymbol{p}+t\boldsymbol{v}$（$t$ 属于 $\mathbb{R}$）的方程.

parametric equation of a plane [一个平面的参数方程] 一个形如 $\boldsymbol{x}=\boldsymbol{p}+s\boldsymbol{u}+t\boldsymbol{v}$（$s,t$ 属于 $\mathbb{R}$）的方程，其中 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是线性无关的向量.

partitioned matrix（of block matrix）[矩阵分划（或分块矩阵）] 一个矩阵，它的元素本身是适当大小的矩阵.

permuted lower triangular matrix [置换下三角形矩阵] 一个矩阵,通过行置换后构成一个下三角形矩阵.

permuted LU factorization [置换 LU 分解] 一个形如 $A=LU$ 的矩阵分解表示，此处 L 是一个行置换后形成的单位下三角形方阵，U 是 A 的一个阶梯形矩阵.

pivot [主元] 一个非零数，或者在主元位置通过行变换用于生成零，或者变成主项为 1，再用于生成零.

pivot column [主元列] 一个包含主元位置的列.

pivot position [主元位置] 矩阵 A 的阶梯矩阵的主元元素所在的位置.

plane through $\boldsymbol{u},\boldsymbol{v}$ and the origin [过 $\boldsymbol{u}$，$\boldsymbol{v}$ 和原点的平面] 一个形如 $\boldsymbol{x}=s\boldsymbol{u}+t\boldsymbol{v}$（$s,t$ 属于 $\mathbb{R}$）的矩阵方程集合，此处 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 线性无关.

polar decomposition（of A）[矩阵 A 的极分解] 一个分解 $A=PQ$，此处 P 是一个 $n\times n$ 半正定矩阵且与 A 的秩相同，Q 是一个 $n\times n$ 正交矩阵.

polygon [多边形] $\mathbb{R}^2$ 中的多面体.

polyhedron [四面体] $\mathbb{R}^3$ 中的多面体.

polytope [多面体] $\mathbb{R}^n$ 中有限点集的凸包（特殊的紧凸集）.

positive combination（of points $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_m$ in $\mathbb{R}^n$）[（$\mathbb{R}^n$ 中的点 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_m$ 的）正组合] 线性组合 $c_1\boldsymbol{v}_1+\cdots+c_m\boldsymbol{v}_m$，这里所有的 $c_i>0$.

positive definite matrix [正定矩阵] 一个对称矩阵 A，使得对所有 $\boldsymbol{x}\neq\mathbf{0}$，$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}>0$ 成立.

positive definite quadratic form [正定二次型] 一个二次型 Q，使得对所有 $\boldsymbol{x}\neq\mathbf{0}$，$Q(\boldsymbol{x})>0$.

positive hull（of a set S）[集合 S 的正包] S 中的点的所有正组合的集合，记为 pos S.

positive semidefinite matrix [半正定矩阵] 一个对称矩阵 A，使得对所有 $\boldsymbol{x}$，$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}\geqslant 0$ 成立.

positive semidefinite quadratic form [半正定二次型] 一个二次型 Q，使得对所有 $\boldsymbol{x}\neq\mathbf{0}$，$Q(\boldsymbol{x})\geqslant 0$.

power method [幂算法] 估计一个方阵严格占优特征值的一个算法.

principal axes（of a quadratic form $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$）[一个二次型 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的主轴] 正交矩阵 P 的单位正交列（这些列是 A 的单位特征向量），使得 $P^{-1}AP$ 是对角阵. 通常 P 的列按 A 的对应特征值大小的递减顺序来排序.

principal components（of the data in a matrix B of observations）[观测矩阵 B 中数据的主成分] B 的样本协方差矩阵 S 的单位特征向量，其特征向量以对应 S 特征值的递减顺序来排列. 如果 B 是平均偏差形式，那么主成分分量是

B^{T} 奇异值分解中的右奇异向量.

probability vector [**概率向量**] 一个 $\mathbb{R}^n$ 中的向量，它的元素非负且和为 1.

product $A\boldsymbol{x}$ [**乘积** $A\boldsymbol{x}$] A 中列的线性组合，利用 $\boldsymbol{x}$ 的对应元素作为权值.

production vector [**生产向量**] 列昂惕夫投入产出模型中的向量，它列出一个经济体系中各部门将要生产的数量.

Profile（of a set S in $\mathbb{R}^n$） [$\mathbb{R}^n$ **中集合** S **的轮廓**] S 的端点的集合.

projection matrix（or orthogonal projection matrix）[**投影矩阵或正交投影矩阵**] 一个对称矩阵 B，使得 $B^2=B$，一个简单例子是 $B=\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}$，此处 $\boldsymbol{v}$ 是一个单位向量.

proper subset of a set S [S **的真子集**] S 的不等于 S 本身的子集.

proper subspace [**真子空间**] 向量空间 V 中，任何一个不是自己本身的子空间.

pseudoinverse（of A）[A **的伪逆**] 矩阵 $VD^{-1}U^{\mathrm{T}}$，当 UDV^{T} 是 A 的一个简化奇异值分解.

Q

QR factorization [QR **分解**] 一个 $m\times n$ 矩阵 A 的分解，其中 A 具有线性无关的列，$A=QR$，此处 Q 是一个 $m\times n$ 矩阵，它的列构成 Col A 的一个单位正交基，R 是一个 $n\times n$ 上三角形可逆矩阵，它的对角线是正元素.

quadratice Bézier curve [**二次贝塞尔曲线**] 一个被描述为如下形式的曲线：$\boldsymbol{g}(t)=(1-t)\boldsymbol{f}_0(t)+t\boldsymbol{f}_1(t)$，其中 $0\leqslant t\leqslant 1$，$\boldsymbol{f}_0(t)=(1-t)\boldsymbol{p}_0+t\boldsymbol{p}_1$，$\boldsymbol{f}_1(t)=(1-t)\boldsymbol{p}_1+t\boldsymbol{p}_2$. 点 $\boldsymbol{p}_0,\boldsymbol{p}_1,\boldsymbol{p}_2$ 称为曲线的控制点.

quadratic form [**二次型**] 一个定义在 $\mathbb{R}^n$ 上的函数 $Q(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$，此处 A 是一个 $n\times n$ 对称矩阵（称为二次型的矩阵）

R

range（of a linear transformation T） [**一个线性变换** T **的值域**] 对于变换 T 定义域中所有 $\boldsymbol{x}$，构成形如 $T(\boldsymbol{x})$ 的所有向量的集合.

rank（of a matrix A）[**矩阵** A **的秩**] A 的列空间的维数，记为 rank A.

Rayleigh quotient [**瑞利商**] $R(\boldsymbol{x})=(\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x})/(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x})$，矩阵 A 的一个特征值估计（A 常为对称矩阵）.

recurrence relation [**递推关系**] 见差分方程.

reduced echelon form（or reduced row echelon form）[**简化的阶梯形式（或简化的行阶梯形式）**] 一个行等价于给定矩阵的简化阶梯矩阵.

reduced echelon matrix [**简化的阶梯矩阵**] 一个阶梯形式的矩形矩阵具有下列附加性质：每一非零行的主元素为 1，每一主元素 1 是它所在列的唯一非零元素.

reduced singular value decomposition [**简化的奇异值分解**] 对秩为 r 的 $m\times n$ 矩阵的一个分解 $A=UDV^{\mathrm{T}}$，此处 U 是具有单位正交列的 $m\times r$ 矩阵，D 是一个 $r\times r$ 对角矩阵，且 D 的对角线上有 r 个非零奇异值，V 是一个具有单位正交列的 $n\times r$ 矩阵.

regression coefficients [**回归系数**] 最小二乘直线 $y=\beta_0+\beta_1 x$ 中的系数 β_0 和 β_1.

Regular solid [**正则立方体**] $\mathbb{R}^3$ 中 5 个可能规则的多面体之一：四面体（4 个等三角面）、立方体（6 个平方面）、八面体（8 个等三角面）、十二面体（12 个相同的 5 角面）、二十面体（20 个相同的三角面）.

regular stochastic matrix [**正则随机矩阵**] 一个随机矩阵 P，使得一些矩阵幂 P^k 仅包含严格正元素.

relative change or relative error（in $\boldsymbol{b}$）[$\boldsymbol{b}$ **中的相对改变或相对误差**] 数量 $\|\Delta\boldsymbol{b}\|/\|\boldsymbol{b}\|$，其中 $\boldsymbol{b}$ 改变为 $\boldsymbol{b}+\Delta\boldsymbol{b}$.

repellor（of a dynamical system in $\mathbb{R}^2$）[$\mathbb{R}^2$ 动力系统的排斥子] 当所有轨迹（除常零序列或常零函数外）都远离 $\mathbf{0}$ 时的原点.

residual vector [残余向量] 一般线性模型 $\boldsymbol{y}=X\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ 中出现的量 $\boldsymbol{\varepsilon}$，即 $\boldsymbol{\varepsilon}=\boldsymbol{y}-X\boldsymbol{\beta}$，关于 y 的观测值和预测值之间的差.

Re $\boldsymbol{x}$ [实部] $\mathbb{R}^n$ 中的向量，它由 $\mathbb{C}^n$ 中的一个向量 $\boldsymbol{x}$ 的元素实部所构成.

right inverse(of A)[A 的右逆] 任何矩形矩阵 C，使得 $AC=I$.

right-multiplication（by A）[用 A 右乘] 用 A 右乘一个矩阵.

right singular vectors（of A）[A 的右奇异向量] 在奇异分解 $A=U\Sigma V^{\mathrm{T}}$ 中，V 的列.

roundoff error [舍入误差] 计算结果被舍入（或截断）到数字的浮点数字存储位数所引起的浮点算术误差.也就是说，一个数（如 1/3）用十进制小数表示为具有有限位数字的浮点数近似结果所产生的误差.

row-column rule [行列法则] 计算乘积 AB 的法则，AB 中 (i,j) 位置的元素是 A 的第 i 行与 B 的第 j 列对应元素乘积之和.

row equivalent（matrices）[行等价矩阵] 两个矩阵通过有限行变换，使得一个矩阵变为另一个矩阵.

row reduction algorithm [行化简算法] 一个系统方法，它利用基本行变换将一个矩阵化简为阶梯形式或简化的阶梯形式.

row replacement [行替换] 一个基本行变换，它将矩阵的一行替换为本行与另一行标量乘法后的和.

row space（of a matrix A）[一个矩阵 A 的行空间] 由 A 的行向量的所有线性组合形成的集合 Row A，同样记作 Col A^{T}.

row sum [行和] 一个矩阵中，一行元素的和.

row vector [行向量] 只有一行的矩阵，或具有几个行的矩阵中的某一行.

row-vector rule for computing $A\boldsymbol{x}$ [计算乘积 $A\boldsymbol{x}$ 的行向量法则] 其中 $A\boldsymbol{x}$ 的第 i 个分量是 A 的第 i 行和向量 $\boldsymbol{x}$ 的对应元素乘积之和.

S

saddle point（of a dynamical system in $\mathbb{R}^2$）[$\mathbb{R}^2$ 中动力系统的鞍点] 当一些轨迹被吸引到 $\mathbf{0}$，另外一些轨迹排斥出 $\mathbf{0}$ 时的原点.

same direction（as a vector $\boldsymbol{v}$）[与向量 $\boldsymbol{v}$ 同一方向] 用正数乘向量 $\boldsymbol{v}$ 得到的向量.

sample mean [样本均值] 一个向量集合 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_N$ 的平均 $\boldsymbol{M}$ 表示为

$$\boldsymbol{M}=(1/N)(\boldsymbol{X}_1+\cdots+\boldsymbol{X}_N)$$

scalar [标量] 一个实数，常用于乘一个向量或者一个矩阵.

scalar multiple of $\boldsymbol{u}$ by c [用标量 c 乘 $\boldsymbol{u}$] 向量 $c\boldsymbol{u}$ 是用标量 c 乘 $\boldsymbol{u}$ 的每一个分量元素来确定的.

scale（a vector）[一个向量的度量] 用一个非零数乘一个向量（或一个矩阵的一行或者一列）.

Schur complement [舒尔补] 一些由一个 2×2 分块矩阵 $A=[A_{ij}]$ 所形成的矩阵.如果 A_{11} 可逆，它的舒尔补是 $A_{22}-A_{21}A_{11}^{-1}A_{12}$.如果 A_{22} 可逆，它的舒尔补是 $A_{11}-A_{12}A_{22}^{-1}A_{21}$.

Schur factorization（of A,for real scalars）[实数矩阵 A 的舒尔分解] 具有 n 个实特征值的 $n\times n$ 矩阵 A 的一个分解 $A=URU^{\mathrm{T}}$，此处 U 是一个 $n\times n$ 正交矩阵，且 R 是一个上三角形矩阵.

set spanned by $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ [由 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 所生成的集合] 集合 $\mathrm{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$.

signal（or discrete-time signal）[信号（或离散时间信号）] 一个两边无穷的数列，$\{y_k\}$；一个定义在整数范围内的函数；它属于向量空间 $\mathbb{S}$.

similar（matrices）[相似矩阵] 矩阵 A 和 B，存在可逆阵 P，使得 $P^{-1}AP=B$ 或 $A=PBP^{-1}$ 成立.

similarity transformation [相似变换] 一个变换将 A 变为 $P^{-1}AP$.

singular (matrix) [奇异矩阵] 没有逆的一个方阵.

singular value decomposition (of an $m\times n$ matrix A) [一个 $m\times n$ 矩阵 A 的奇异值分解] $A=U\Sigma V^{\mathrm{T}}$, 此处 U 是一个 $m\times m$ 正交矩阵, V 是一个 $n\times n$ 正交矩阵, Σ 是一个 $m\times n$ 矩阵且非零元素位于主对角线上(大小以递减顺序排列), 零位于其他位置. 如果 rank $A=r$, 那么 Σ 恰好具有 r 个正元素 (A 的非零奇异值) 在其对角线上.

singular values (of A) [A 的奇异值] 矩阵 $A^{\mathrm{T}}A$ 特征值的正的平方根, 其大小按递减顺序排列.

size (of a matrix) [矩阵的大小] 写成 $m\times n$ 的形式的两个数, 给出一个矩阵中行的数目 m 和列的数目 n.

solution (of a linear system involving variables $(x_1,\cdots,x_n)$) [关于变量 $x_1,\cdots,x_n$ 的线性方程组的解] 一列数 $(s_1,s_2,\cdots,s_n)$ 使得当值 $s_1,\cdots,s_n$ 分别代替 $x_1,\cdots,x_n$ 后, 线性方程组中每一个方程都成立.

solution set [解集] 线性方程所有解的集合.

Span $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ [$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 所生成的子空间] $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 的所有线性组合的集合, 也是由 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 所生成 (或张成) 的子空间.

spanning set (for a subspace H) [一个子空间 H 的生成集] H 中的任何集合 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 使得 $H=\mathrm{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$.

spectral decomposition (of A) [A 的谱分解] 表达式 $A=\lambda_1\boldsymbol{u}_1\boldsymbol{u}_1^{\mathrm{T}}+\cdots+\lambda_n\boldsymbol{u}_n\boldsymbol{u}_n^{\mathrm{T}}$. 此处, $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n\}$ 是 A 的特征向量的一个单位正交基, 且 $\lambda_1,\cdots,\lambda_n$ 是 A 的对应特征值.

spiral point (of a dynamical system in $\mathbb{R}^2$) [$\mathbb{R}^2$ 中动力系统的螺线极点] 当轨迹绕 0 螺旋时的原点.

stage-matrix model [分级矩阵模型] 一个差分方程 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$, 此处 $\boldsymbol{x}_k$ 列出时间 k 时, 全体人口中女性的数目, 其中女性根据不同发展阶段来分类 (如青少年、接近成年、成年).

standard basis [标准基] 由 $n\times n$ 单位矩阵的列所构成的 $\mathbb{R}^n$ 的基, $\mathcal{E}=\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$, 或 $\mathbb{P}_n$ 的基 $\{1,t,\cdots,t^n\}$.

standard matrix (for a linear transformation T) [一个线性变换 T 的标准矩阵] 矩阵 A 使得 $T(\boldsymbol{x})=A\boldsymbol{x}$, 对 T 的定义域中所有 $\boldsymbol{x}$ 成立.

standard position [标准位置] 当 A 是一个对角矩阵时, 方程 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=c$ 的图形所处的位置.

state vector [状态向量] 一个概率向量, 一般地, 一个向量表示一个物理系统的"状态", 常常与一个差分方程 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 联系在一起.

steady-state vector (for a stochastic matrix P) [一个随机矩阵 P 的稳态向量] 存在一个概率向量 $\boldsymbol{q}$ 使得 $P\boldsymbol{q}=\boldsymbol{q}$ 成立.

stiffness matrix [刚性矩阵] 一个弹性矩阵的逆, 刚性矩阵的第 j 列给出相应的负载施加于一个弹性梁的特定点, 使得在梁的第 j 点产生一个单位弯曲.

stochastic matrix [随机矩阵] 一个方阵, 其列是随机向量.

strictly dominant eigenvalue [严格占优特征值 (主特征值)] 矩阵 A 的一个特征根 λ_1, 对所有 A 的其他特征值 λ_k 具有性质 $|\lambda_1|>|\lambda_k|$.

submatrix (of A) [A 的子矩阵] 任何由划去 A 的一些行和 A 的一些列所得到的矩阵, A 自身也是.

subspace [子空间] 由 V 中一些向量构成的子集 H, 使得 H 具有性质: (1) V 中的零向量在 H 中; (2) H 对向量加法封闭; (3) H 对标量乘法封闭.

supporting hyperplane (to a compact convex set S in $\mathbb{R}^n$) [支撑超平面 ($\mathbb{R}^n$ 的紧致凸集 S)

一个超平面 $H=[f:d]$，使得 $H\cap S\neq\varnothing$，并且对 S 中所有的 x，有 $f(x)\leqslant d$ 或 $f(x)\geqslant d$ 成立.

symmetric matrix [对称矩阵] 一个矩阵使得 $A^{\mathrm{T}}=A$.

system of linear equations（or a linear system）[线性方程组或一个线性系统] 一个或多个与相同变量集如 $x_1,\cdots,x_n$ 有关的线性方程的集合.

T

tetrahedron [四面体] 由 4 个相等三角平面界定的三维立方体，这里每三个平面共一个顶点.

total variance [方差总和] 一个观测矩阵的协方差矩阵的迹.

trace（of a square matrix A）[一个方阵 A 的迹] A 中对角线元素之和，记为 $\operatorname{tr} A$.

trajectory [轨迹] 一个动力系统 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 解的图像 $\{\boldsymbol{x}_0,\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots\}$. 常与一条曲线联系起来，使得轨迹容易观察. 也是 $t\geqslant 0$ 时，$\boldsymbol{x}(t)$ 的图像，其中 $\boldsymbol{x}(t)$ 是微分方程 $\boldsymbol{x}'(t)=A\boldsymbol{x}(t)$ 的解.

transfer matrix [转换矩阵] 与一个电路有关的矩阵，具有输入和输出端，使得输出向量是 A 与输入向量乘积.

transformation（or function or mapping）T from $\mathbb{R}^n$ to $\mathbb{R}^m$ [从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的变换 T（或函数、映射）] 一个法则，它指定 $\mathbb{R}^n$ 中每一个向量 $\boldsymbol{x}$ 对应 $\mathbb{R}^m$ 中唯一的向量 $T(\boldsymbol{x})$. 记号：$T:\mathbb{R}^n\to\mathbb{R}^m$，也是 $T:V\to W$ 表示一个法则，它指定 V 中一个 $\boldsymbol{x}$ 对应 W 中唯一的向量 $T(\boldsymbol{x})$.

translation（by a vector $\boldsymbol{p}$）[平移一个向量 $\boldsymbol{p}$ 的变换] 一个向量加 $\boldsymbol{p}$ 后的变换，或对给定集合中任一向量加 $\boldsymbol{p}$ 的变换.

transpose（of A）[A 的转置] 一个 $n\times m$ 矩阵 A^{T}，它的列对应 $m\times n$ 矩阵 A 的行.

trend analysis [趋势分析] 通过计算在有限点集处的内积，用正交多项式拟合数据.

triangle inequality [三角不等式] 对所有 $\boldsymbol{u}$ 和 $\boldsymbol{v}$，$\|\boldsymbol{u}+\boldsymbol{v}\|\leqslant\|\boldsymbol{u}\|+\|\boldsymbol{v}\|$ 成立.

triangular matrix [三角矩阵] 一个矩阵 A，或者对角线上方的元素为 0，或者对角线下方的元素为 0.

trigonometric polynomial [三角多项式] 由常数函数 1 和正弦函数和余弦函数，例如 $\cos nt$ 和 $\sin nt$ 的线性组合所构成.

trivial solution [平凡解] 齐次线性方程组 $A\boldsymbol{x}=\boldsymbol{0}$ 的解，$\boldsymbol{x}=\boldsymbol{0}$.

U

uncorrelated variables [不相关变量] 任意两个变量 x_i 和 $x_j(i\neq j)$，它们取遍一个观测矩阵中观测向量的 i 坐标和 j 坐标，且协方差 s_{ij} 是零.

underdetermined system [不确定系统] 方程个数少于未知数个数的线性方程组.

uniqueness question [唯一性问题] “方程组是否有解？它的解是否唯一？即只有一个解吗？”

unit consumption vector [单位消费向量] 列昂惕夫投入产出模型中的列向量，它给出每一单位的输出所需的每一个部门的输入；是消费矩阵的一个列.

unit lower triangular matrix [单位下三角形矩阵] 一个下三角形方阵，其主对角线上的元素是 1.

unit vector [单位向量] 一个满足 $\|\boldsymbol{v}\|=1$ 的向量 $\boldsymbol{v}$.

upper triangular matrix [上三角形矩阵] 一个矩阵 U（不一定方阵），其零元素位于对角线上元素 $u_{11},u_{22},\cdots$ 之下方.

V

Vandermonde matrix [范德蒙德矩阵] 一个矩阵 V 或它的转置，V 的形式是

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix}$$

variance (of a variable x_j) [一个变量 x_j 的方差] 一个观测矩阵的协方差矩阵 S 的对角线上元素 s_{jj}，此处 x_j 取遍观测向量的 j 个坐标.

vector [向量] 一列数；只有一列的矩阵，一般指一个向量空间中的任一元素.

vector addition [向量加法] 通过对应元素相加得到向量相加.

vector equation [向量方程] 含未知权值的向量线性组合所得到的方程.

vector space [向量空间] 以向量为对象所形成的集合，且定义被称为加法和标量乘法的两种运算，必须满足十个公理，见 4.1 节的第一个定义.

vector subtraction [向量减法] 计算 $\boldsymbol{u}+(-1)\boldsymbol{v}$ 且将其写成 $\boldsymbol{u}-\boldsymbol{v}$.

W

weighted least squares [带权的最小二乘] 带有权值内积所对应的最小二乘问题，例如

$$<\boldsymbol{x},\boldsymbol{y}> = w_1^2 x_1 y_1 + \cdots + w_n^2 x_n y_n$$

weights [权值] 一个线性组合中所用的数.

Z

zero subspace [零空间] 仅包含一个零向量的空间 $\{\boldsymbol{0}\}$.

zero vector [零向量] 唯一的向量，记为 $\boldsymbol{0}$，使得 $\boldsymbol{u}+\boldsymbol{0}=\boldsymbol{u}$ 对所有 $\boldsymbol{u}$ 成立，在 $\mathbb{R}^n$ 中，$\boldsymbol{0}$ 是一个所有元素为零的向量.

奇数习题答案

第 1 章

1.1

1. 解是$(x_1,x_2)=(-8,3)$，或简单写成$(-8,3)$.

3. （2,1）

5. 替换第 2 行为第 2 行加–4 乘以第 3 行，然后替换第 1 行为第 1 行加 3 乘以第 3 行.

7. 解集为空集.

9. （16,21,14,4）

11. 不相容

13. （5,3,–1）

15. 不相容

17. 计算表明方程组不相容，因此这三条线没有交点.

19. $h\neq 2$

21. 所有 h

23. 标记一个命题为真当且仅当这个命题总是真的. 只给出答案会达不到判断题的意义，即帮助你学会仔细阅读教材是重要的. “学习指导”告诉你在哪里可以找到答案，但你不应该在没有尝试自己寻找答案之前就去查看“学习指导”.

25. $k-2g+h=0$

27. 行化简$\begin{bmatrix} a & b & f \\ c & d & g \end{bmatrix}$为$\begin{bmatrix} a & b & f \\ 0 & d-b\left(\dfrac{c}{a}\right) & g-f\left(\dfrac{c}{a}\right) \end{bmatrix}$后显示$d-b\left(\frac{c}{a}\right)$必定不为零，因为 f 和 g 是任意值. 否则，对 f 和 g 的某种取值，第 2 行对应于形式为$0=q$的方程，其中 q 不为零. 因此$ad\neq bc$.

29. 交换第 1 行和第 3 行；交换第 1 行和第 3 行.

31. 用第 3 行加–4 乘第 1 行替换第 3 行；用第 3 行加 4 乘第 1 行替换第 3 行.

33. 回顾练习题 1，写出一个解. “学习指导”有一个解.

1.2

1. 简化阶梯形：a,b；阶梯形：d；非阶梯形：c.

3. $\begin{bmatrix} 1 & 2 & 0 & -8 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$，主元列是第 1,3 列：$\begin{bmatrix} 1 & 2 & 4 & 8 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 9 & 12 \end{bmatrix}$.

5. $\begin{bmatrix} ■ & * \\ 0 & ■ \end{bmatrix},\begin{bmatrix} ■ & * \\ 0 & 0 \end{bmatrix},\begin{bmatrix} 0 & ■ \\ 0 & 0 \end{bmatrix}$

7. $\begin{cases} x_1=-5-3x_2 \\ x_2\text{是自由变量} \\ x_3=3 \end{cases}$

9. $\begin{cases} x_1=3+2x_3 \\ x_2=3+2x_3 \\ x_3\text{是自由变量} \end{cases}$

11. $\begin{cases} x_1=\dfrac{2}{3}x_2-\dfrac{4}{3}x_3 \\ x_2\text{是自由变量} \\ x_3\text{是自由变量} \end{cases}$

13. $\begin{cases} x_1=5+3x_5 \\ x_2=1+4x_5 \\ x_3\text{是自由变量} \\ x_4=4-9x_5 \\ x_5\text{是自由变量} \end{cases}$

15. a. 相容，有多解　b. 相容，有多解

17. 所有 h

19. a. 当 $h=2,k\neq 8$ 时不相容

    b. 当 $h\neq 2$ 时有唯一解

    c. 当 $h=2,k=8$ 时有多解

21. 仔细阅读课本，查阅“学习指导”之前写出你的答案，记住，一个论断是正确的，当且仅当所有情形论断都正确.

23. 因为有 4 个主元（系数矩阵的每列有一个主元）. 讨论的矩阵必可简化为

$$\begin{matrix} 1 & 0 & 0 & 0 & a \\ 0 & 1 & 0 & 0 & b \\ 0 & 0 & 1 & 0 & c \\ 0 & 0 & 0 & 1 & d \end{matrix}$$

因此，

$$\begin{aligned} x_1 &= a \\ x_2 &= b \\ x_3 &= c \\ x_4 &= d \end{aligned}$$

无论 a,b,c 和 d 取何值，解都存在且唯一.

25. 如果系数矩阵中的每一行有一个主元位置，那

么在底行有一个主元位置，从而在增广列没有主元位置，所以，根据定理 2，方程组是相容的.

27. 如果方程组是相容的，则解是唯一的当且仅当系数矩阵的每一列都是主元列，否则，方程组有无穷多解.

29. 一个不确定的方程组中，未知变量的个数总多于方程的个数，基本变量的个数不多于方程的个数，所以，至少有一个自由变量，这个变量可以被指定无穷多个值. 如果方程组是相容的，自由变量的每一个不同值会产生一个不同解.

31. 是的，一个线性方程组的方程个数多于未知变量的个数，它可以是相容的. 下面的方程组有一个解 $(x_1 = x_2 = 1)$：

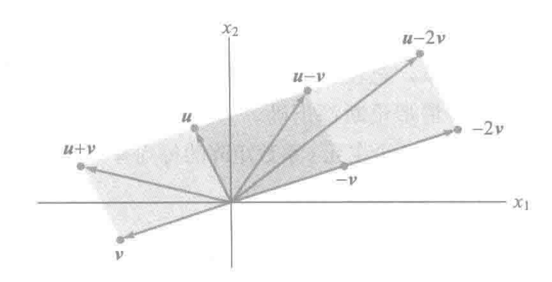
$$\begin{aligned} x_1 + \ \ x_2 &= 2 \\ x_1 - \ \ x_2 &= 0 \\ 3x_1 + 2x_2 &= 5 \end{aligned}$$

33. [M] $p(t) = 1 + 3t + 2t^2$

1.3

1. $\begin{bmatrix} -4 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 5 \\ 4 \end{bmatrix}$

3.



5. $x_1\begin{bmatrix} 3 \\ -2 \\ 8 \end{bmatrix} + x_2\begin{bmatrix} 5 \\ 0 \\ -9 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 8 \end{bmatrix}$,

$$\begin{bmatrix} 3x_1 \\ -2x_1 \\ 8x_1 \end{bmatrix} + \begin{bmatrix} 5x_2 \\ 0 \\ -9x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 8 \end{bmatrix}, \quad \begin{bmatrix} 3x_1+5x_2 \\ -2x_1 \\ 8x_1-9x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 8 \end{bmatrix}$$

$$\begin{aligned} 3x_1+5x_2 &= 2 \\ -2x_1 \qquad &= -3 \\ 8x_1-9x_2 &= 8 \end{aligned}$$

通常不写中间过程.

7. $\boldsymbol{a} = \boldsymbol{u} - 2\boldsymbol{v}$, $\boldsymbol{b} = 2\boldsymbol{u} - 2\boldsymbol{v}$, $\boldsymbol{c} = 2\boldsymbol{u} - 3.5\boldsymbol{v}$, $\boldsymbol{d} = 3\boldsymbol{u} - 4\boldsymbol{v}$

是的，$\mathbb{R}^2$ 中的每个向量是 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的线性组合.

9. $x_1\begin{bmatrix} 0 \\ 4 \\ -1 \end{bmatrix} + x_2\begin{bmatrix} 1 \\ 6 \\ 3 \end{bmatrix} + x_3\begin{bmatrix} 5 \\ -1 \\ -8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

11. 否，$\boldsymbol{b}$ 不是 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ 的线性组合

13. 否，$\boldsymbol{b}$ 不是 A 的列的线性组合.

15. $h = 3$

17. 当然，非整数权值也可以接受，但一些简单选择是 $0 \cdot \boldsymbol{v}_1 + 0 \cdot \boldsymbol{v}_2 = \boldsymbol{0}$，且

$$1 \cdot \boldsymbol{v}_1 + 0 \cdot \boldsymbol{v}_2 = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}, \ 0 \cdot \boldsymbol{v}_1 + 1 \cdot \boldsymbol{v}_2 = \begin{bmatrix} -4 \\ 0 \\ 1 \end{bmatrix}$$

$$1 \cdot \boldsymbol{v}_1 + 1 \cdot \boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix}, \ 1 \cdot \boldsymbol{v}_1 - 1 \cdot \boldsymbol{v}_2 = \begin{bmatrix} 7 \\ 1 \\ 1 \end{bmatrix}$$

19. $\mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ 是在通过 $\boldsymbol{v}_1$ 和原点的直线上的点集，因为 $\boldsymbol{v}_2$ 是 $\boldsymbol{v}_1$ 的倍数.

21. 提示：证明 $\begin{bmatrix} 2 & 2 & h \\ -1 & 1 & k \end{bmatrix}$ 对所有 h 和 k 是相容的，解释这个计算能说明 $\mathrm{Span}\{\boldsymbol{u}, \boldsymbol{v}\}$ 是什么.

23. 在你查阅“学习指导”之前，先仔细阅读整节内容. 特别注意定义和定理的叙述，和在其前后的言论.

25. a. 否，3　　b. 是，无穷多

c. $\boldsymbol{a}_1 = 1 \cdot \boldsymbol{a}_1 + 0 \cdot \boldsymbol{a}_2 + 0 \cdot \boldsymbol{a}_3$

27. a. $5\boldsymbol{v}_1$ 是 5 天工作 1#矿的产出.

b. 总的产出是 $x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2$，所以 x_1 和 x_2 应该满足 $x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 = \begin{bmatrix} 240 \\ 2\,824 \end{bmatrix}$.

c. [M]1.73 天生产 1#矿和 4.70 天生产 2#矿.

29. $(17/14, -34/14, 16/14) = (17/14, -17/7, 8/7)$

31. a. $\begin{bmatrix} 10/3 \\ 2 \end{bmatrix}$

b. 在（0,1）点加 3.5 克，在（8,1）点加 0.5 克，在（2,4）点加 2 克.

33. 复习练习题 1 再写出解答，“学习指导”有参考解答.

1.4

1. 乘积无定义，因为 3×2 矩阵的列数为 2，而向量的元素个数是 3，两者不匹配.

3. a. $Ax=\begin{bmatrix}1&2\\-3&1\\1&6\end{bmatrix}\begin{bmatrix}-2\\3\end{bmatrix}=-2\cdot\begin{bmatrix}1\\-3\\1\end{bmatrix}+3\cdot\begin{bmatrix}2\\1\\6\end{bmatrix}$

$$=\begin{bmatrix}-2\\6\\6\end{bmatrix}+\begin{bmatrix}-6\\3\\18\end{bmatrix}=\begin{bmatrix}4\\9\\16\end{bmatrix}$$

b. $Ax=\begin{bmatrix}1&2\\-3&1\\1&6\end{bmatrix}\begin{bmatrix}-2\\3\end{bmatrix}=\begin{bmatrix}1\cdot(-2)+2\cdot(3)\\(-3)\cdot(-2)+1\cdot(3)\\1\cdot(-2)+6\cdot(3)\end{bmatrix}=\begin{bmatrix}4\\9\\16\end{bmatrix}$.

在这里以及习题 4~6 展示你的工作，以后的习题计算过程用心算.

5. $2\cdot\begin{bmatrix}1\\-2\end{bmatrix}-1\cdot\begin{bmatrix}2\\-3\end{bmatrix}+1\cdot\begin{bmatrix}-3\\1\end{bmatrix}-1\cdot\begin{bmatrix}1\\-1\end{bmatrix}=\begin{bmatrix}-4\\1\end{bmatrix}$

7. $\begin{bmatrix}4&-5&7\\-1&3&-8\\7&-5&0\\-4&1&2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}6\\-8\\0\\-7\end{bmatrix}$

9. $x_1\begin{bmatrix}5\\0\end{bmatrix}+x_2\begin{bmatrix}1\\2\end{bmatrix}+x_3\begin{bmatrix}-3\\4\end{bmatrix}=\begin{bmatrix}8\\0\end{bmatrix}$，$\begin{bmatrix}5&1&-3\\0&2&4\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}8\\0\end{bmatrix}$

11. $\begin{bmatrix}1&3&-4&-2\\1&5&2&4\\-3&-7&6&12\end{bmatrix}$，$x=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}-11\\3\\0\end{bmatrix}$

13. 是（验证你的答案）.

15. 当 $3b_1+b_2$ 非零时，方程 $Ax=b$ 不相容.（给出理由.）所有使方程 $Ax=b$ 相容的 b 组成的集合，是一条通过原点的直线 $b_2=-3b_1$.

17. 只有 3 行含有主元. 由定理 4，方程 $Ax=b$ 不是对 $\mathbb{R}^4$ 中的每个 b 都有解.

19. 习题 17 说明定理 4 命题（d）不真，因此不是 $\mathbb{R}^4$ 中的所有向量都可以写成 A 的列的线性组合有解，即 A 的列不能生成 $\mathbb{R}^4$.

21. 矩阵 $[v_1\ \ v_2\ \ v_3]$ 不是每一行都有主元，因此由定理 4，矩阵的列不能生成 $\mathbb{R}^4$，即 $\{v_1,v_2,v_3\}$ 不能生成 $\mathbb{R}^4$.

23. 仔细阅读教材，在查阅“学习指导”之前先判断习题中每个命题的真假. 习题 29~30 中的许多命题是这种形式的蕴涵式：“如果 <命题 1>，则<命题 2>”或等价地，“有<命题 2>，如果<命题 1>成立.”当<命题 1>为真时，如果在任何情形下<命题 2>都成立，则将这样的蕴涵式标记为真.

25. $c_1=-3, c_2=-1, c_3=2$

27. 矩阵方程可以写为
$c_1v_1+c_2v_2+c_3v_3+c_4v_4+c_5v_5=v_6$，其中 $c_1=-3$，$c_2=1$，$c_3=2$，$c_4=-1$，$c_5=2$，并且
$v_1=\begin{bmatrix}-3\\5\end{bmatrix}$，$v_2=\begin{bmatrix}5\\8\end{bmatrix}$，$v_3=\begin{bmatrix}-4\\1\end{bmatrix}$，
$v_4=\begin{bmatrix}9\\-2\end{bmatrix}$，$v_5=\begin{bmatrix}7\\-4\end{bmatrix}$，$v_6=\begin{bmatrix}11\\-11\end{bmatrix}$，

29. 提示：从有 3 个主元位置的 3×3 阶梯形矩阵 B 开始寻找.

31. 查阅“学习指导”之前写出你的解答.

33. 提示：A 有多少主元列？为什么？

35. 假设 y 和 z 满足 $Ay=z$，则 $5z=5Ay$. 由定理 5（b），$5Ay=A(5y)$. 因此 $5z=A(5y)$，这说明 $5y$ 是 $Ax=5z$ 的一个解. 从而方程 $Ax=5z$ 是相容的.

37. [M] 列不生成 $\mathbb{R}^4$.

39. [M] 列生成 $\mathbb{R}^4$.

41. [M] 在习题 39 中划掉矩阵的第 4 列. 也可以不划掉矩阵的第 4 列，而是划掉矩阵的第 3 列.

1.5

1. 方程组有非平凡解，因为有自由变量 x_3.

3. 方程组有非平凡解，因为有自由变量 x_3.

5. $x=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=x_3\begin{bmatrix}-1\\-1\\1\end{bmatrix}$

7. $x=\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=x_3\begin{bmatrix}-9\\4\\1\\0\end{bmatrix}+x_4\begin{bmatrix}8\\-5\\0\\1\end{bmatrix}$

9. $x = x_2\begin{bmatrix}2\\1\\0\end{bmatrix}$

11. 提示：由阶梯形推出的线性方程组是

$$\begin{aligned} x_1 - 4x_2 \quad\quad + 5x_6 &= 0 \\ x_3 \quad - x_6 &= 0 \\ x_5 - 4x_6 &= 0 \\ 0 &= 0 \end{aligned}.$$

基本变量是 x_1, x_3, x_5，其他变量是自由变量．“学习指导”讨论了两种这类型题目常犯的错误．

13. $x = \begin{bmatrix}5\\-2\\0\end{bmatrix} + x_3\begin{bmatrix}4\\-7\\1\end{bmatrix} = p + x_3 q$．在几何上，解集是通过 $\begin{bmatrix}5\\-2\\0\end{bmatrix}$ 且平行于 $\begin{bmatrix}4\\-7\\1\end{bmatrix}$ 的直线．

15. 令 $u = \begin{bmatrix}-5\\1\\0\end{bmatrix}, v = \begin{bmatrix}3\\0\\1\end{bmatrix}, p = \begin{bmatrix}-2\\0\\0\end{bmatrix}$，齐次线性方程组的解是 $x = x_2 u + x_3 v$，由 u 和 v 所生成的通过原点的平面．非齐次方程的解集是 $x = p + x_2 u + x_3 v$，此平面通过 p，平行于齐次线性方程组解集的平面．

17. $x = \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}8\\-4\\0\end{bmatrix} + x_3\begin{bmatrix}-1\\-1\\1\end{bmatrix}$．解集是通过 $\begin{bmatrix}8\\-4\\0\end{bmatrix}$ 的直线，平行于习题 5 中齐次线性方程组解集的直线．

19. $x = a + tb$，其中 t 为参数，或者

$$x = \begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}-2\\0\end{bmatrix} + t\begin{bmatrix}-5\\3\end{bmatrix} \text{或} \begin{cases}x_1 = -2 - 5t\\ x_2 = 3t\end{cases}$$

21. $x = p + t(q - p) = \begin{bmatrix}3\\-3\end{bmatrix} + t\begin{bmatrix}1\\4\end{bmatrix}$

23. 非常重要的是仔细阅读课本，然后写出你的答案．完成之后，如果需要，用“学习指导”校对答案．

25. a. $Aw = A(p + v_h) = Ap + Av_h = b + 0 = b$

 b. $Av_h = A(w - p) = Aw - Ap = b - b = 0$

27. （利用定理 6 进行几何讨论）因为方程 $Ax = b$ 相容，根据定理 6，它的解集可以通过平移 $Ax = 0$ 的解集获得．因此，$Ax = b$ 的解集是一个单向量，当且仅当 $Ax = 0$ 的解集是一个单向量，而这当且仅当 $Ax = 0$ 只有平凡解才发生．（使用自由变量证明）若 $Ax = b$ 有一个解，则解是唯一的，当且仅当相应方程组中没有自由变量，即当且仅当 A 的每列都是主元列．这当且仅当 $Ax = 0$ 仅有平凡解才发生．

29. a. 当 A 是 4×4 矩阵，有 3 个主元位置，方程 $Ax = 0$ 有一个自由变量，因此没有非平凡解．

 b. 有 3 个主元位置时，A 在每一行没有一个主元，由 1.4 节定理 4，方程 $Ax = b$ 对每个可能的 b 都没有解．“可能”的意思是此处考虑的向量是 $\mathbb{R}^4$ 中的向量，因为 A 有 4 行．

31. a. 当 A 是 3×2 矩阵，有 2 个主元位置，每一列是主元列，方程 $Ax = 0$ 没有自由变量，因此没有非平凡解．

 b. 有 2 个主元位置和三行时，A 不可能在每一行有一个主元，由 1.4 节定理 4，方程 $Ax = b$ 不可以对每个 $\mathbb{R}^3$ 中可能的 b 都有一个解．

33. 你的例子中应该有每行的元素之和等于零. 为什么?

35. 一个答案是 $x = \begin{bmatrix}3\\-1\end{bmatrix}$．

37. 一个答案是 $A = \begin{bmatrix}1 & -4\\1 & -4\end{bmatrix}$．“学习指导”给出如何分析问题以构造 A．如果 b 不是 A 的第一列的数倍，则 $Ax = b$ 的解集为空，因此不可能平移 $Ax = 0$ 的解集来得到 $Ax = b$ 的解．这并不与定理 6 相矛盾，因为定理 6 仅当方程 $Ax = b$ 有非空的解集时才成立．

39. 如果 $Av=0$，$Aw=0$，则因为由 1.4 节定理 5（a）有 $A(v+w)=Av+Aw$．现在令 c 和 d 是数，利用定理 5，$A(cv+dw)=A(cv)+A(dw)=cAv+dAw=c0+d0=0$．

1.6

1. 一通解是 $p_G = 0.875p_S$，此处 p_S 是自由变量，一个平衡的解是 $p_S = 1000$ 和 $p_G = 875$．使用分数，一般解可写成 $p_G = (7/8)p_S$，一个自然的价格取值是 $p_S = 80$，$p_G = 70$．只有价格的比例是重要的，经济平衡不会受一个价格变化的影响．

3. a.

产出分配

燃料能源	加工	服务		
产出 ↓	↓	↓	投入	购买部门
0.1	0.1	0.2	→	燃料能源
0.8	0.1	0.4	→	加工
0.1	0.8	0.4	→	服务

b. $\begin{bmatrix} 0.9 & -0.1 & -0.2 & 0 \\ -0.8 & 0.9 & -0.4 & 0 \\ -0.1 & -0.8 & 0.6 & 0 \end{bmatrix}$

c. [M] $p_{\text{F\&P}} \approx 30, p_{\text{M}} \approx 71, p_{\text{S}} = 100$.

5. a.

产出分配

农业	加工	服务	运输		
产出 ↓	↓	↓	↓	投入	购买部门
0.20	0.35	0.10	0.20	→	农业
0.20	0.10	0.20	0.30	→	加工
0.30	0.35	0.50	0.20	→	服务
0.30	0.20	0.20	0.30	→	运输

b. 一个解是 p_{A}=7.99，p_{M}=8.36，p_{S}=14.65，p_{T}=10.00.

c.

产出分配

农业	加工	服务	运输		
产出 ↓	↓	↓	↓	投入	购买部门
0.20	0.35	0.10	0.20	→	农业
0.10	0.10	0.20	0.30	→	加工
0.40	0.35	0.50	0.20	→	服务
0.30	0.20	0.20	0.30	→	运输

d. 一个解是 p_{A}=7.81，p_{M}=7.67，p_{S}=15.62，p_{T}=10.00.

服务部门最受益.

7. $3NaHCO_3 + H_3C_6H_5O_7 \rightarrow Na_3C_6H_5O_7 + 3H_2O + 3CO_2$

9. $B_2S_3 + 6H_2O \rightarrow 2H_3BO_3 + 3H_2S$

11. [M] $16MnS + 13As_2Cr_{10}O_{35} + 374H_2SO_4 \rightarrow 16HMnO_4 + 26AsH_3 + 130CrS_3O_{12} + 327H_2O$

13. a. $\begin{cases} x_1 = x_3 - 40 \\ x_2 = x_3 + 10 \\ x_3\text{是自由变量} \\ x_4 = x_6 + 50 \\ x_5 = x_6 + 60 \\ x_6\text{是自由变量} \end{cases}$ b. $\begin{cases} x_2 = 50 \\ x_3 = 40 \\ x_4 = 50 \\ x_5 = 60 \end{cases}$

15. $\begin{cases} x_1 = 60 + x_6 \\ x_2 = -10 + x_6 \\ x_3 = 90 + x_6 \\ x_4 = x_6 \\ x_5 = 80 + x_6 \\ x_6\text{是自由变量} \end{cases}$

为了流量为非负，$x_6 \geqslant 10$

1.7

对习题 1~22 验证你的答案.

1. 线性无关 3. 线性相关

5. 线性无关 7. 线性相关

9. a. 没有 h b. 所有 h

11. $h = -4$ 13. 所有 h

15. 线性相关 17. 线性相关

19. 线性无关

21. 如果在思考判断题之前查阅“学习指导”，会失去做这类题目的意义.

23. $\begin{bmatrix} \blacksquare & * \\ 0 & 0 \end{bmatrix}$，$\begin{bmatrix} 0 & \blacksquare \\ 0 & 0 \end{bmatrix}$，$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

25. $\begin{bmatrix} \blacksquare & * \\ 0 & \blacksquare \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$ 或 $\begin{bmatrix} 0 & \blacksquare \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

27. 6×4 矩阵 A 的所有 4 列一定是主元列. 否则方程 $A\boldsymbol{x} = \boldsymbol{0}$ 有自由变量，此时，A 的列线性相关.

29. A:任何有一列是另外一列倍数的 3×2 矩阵.
B:任何有两个非零列的 3×2 矩阵，但两列不是相互有倍数关系，此时，这两列线性无关，因此方程 $A\boldsymbol{x} = \boldsymbol{0}$ 只有平凡解.

31. $x=\begin{bmatrix}1\\1\\-1\end{bmatrix}$

33. 由定理 7 可知，真. （“学习指导”补充了另外的验证.）

35. 由定理 9 可知，真.

37. 正确，v_1,v_2,v_3 之间的线性相关，通过选取 v_4 的权值为零可以扩展为 v_1,v_2,v_3,v_4 之间的线性相关.

39. 你有能力在没有帮助的条件下做这个重要的题目. 在查阅“学习指导”之前写出你的解答.

41. [M]利用 A 的主元列，$B=\begin{bmatrix}3&-4&7\\-5&-3&-11\\4&3&2\\8&-7&4\end{bmatrix}$，可能有其他选择.

43. [M] A 中不属于 B 的每一个列，属于 B 的列所生成的集合.

1.8

1. $\begin{bmatrix}2\\-6\end{bmatrix},\begin{bmatrix}2a\\2b\end{bmatrix}$

3. $x=\begin{bmatrix}7\\6\\3\end{bmatrix}$，唯一解

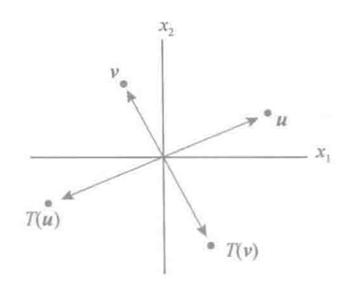
5. $x=\begin{bmatrix}3\\1\\0\end{bmatrix}$，不唯一

7. $a=5,\ b=6$

9. $x=x_3\begin{bmatrix}4\\3\\1\\0\end{bmatrix}$

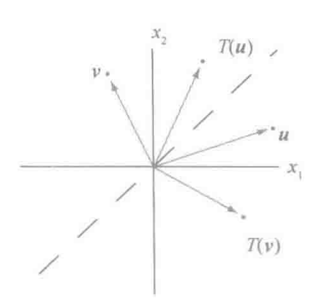
11. 是，因为 $[A\ \ b]$ 所表示的方程组是相容的.

13.



通过原点的对称

15.



通过直线 $x_1=x_2$ 的对称

17. $\begin{bmatrix}8\\2\end{bmatrix},\begin{bmatrix}-3\\9\end{bmatrix},\begin{bmatrix}5\\11\end{bmatrix}$

19. $\begin{bmatrix}13\\7\end{bmatrix},\begin{bmatrix}2x_1-x_2\\5x_1+6x_2\end{bmatrix}$

21. 仔细阅读课本，在查阅“学习指导”之前写出你的答案，注意练习 21（e）是下列形式的句子，

“〈命题 1〉当且仅当〈命题 2〉”

标出这样的句子是真的，如果〈命题 1〉真时，在任何情况下〈命题 2〉是真的，并且如果〈命题 2〉真时，在任何情况下〈命题 1〉是真的.

23. a. 当 $b=0, f(x)=mx$，对任何 $\mathbb{R}$ 中的 x,y 和任何数 c,d.

$$f(cx+dy)=m(cx+dy)=mcx+mdy$$
$$=c(mx)+d(my)=c\cdot f(x)+d\cdot f(y)$$

这说明 f 是线性的.

b. 当 $f(x)=mx+b$ 时，其中 b 非零，$f(0)=m(0)+b=b\neq 0$.

c. 在微积分中，f 称为线性函数，因为 f 的图像是一直线.

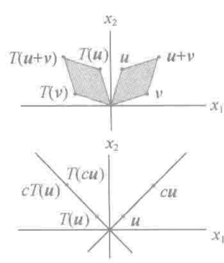
25. 提示：利用直线的参数方程表示直线的像（即直线上的点的像组成的集合）.

27. 对某些 s,t，平面 P 上的任何点 x 满足参数方程 $x=su+tv$. 由线性，$T(x)$的象满足参数方程

$$T(x)=sT(u)+tT(v), s,t\in\mathbb{R}$$

象集合由 Span$\{T(u),T(v)\}$张成. 若 $T(u)$与 $T(v)$是线性无关的，Span$\{T(u),T(v)\}$是过 $T(u)$与 $T(v)$与原点的平面. 若 $T(u)$与 $T(v)$是线性相关的且不都为零，则 Span$\{T(u),T(v)\}$是过原点的直线. 若 $T(u)=T(v)=0$，则 Span$\{T(u),T(v)\}$是$\{0\}$.

29.



31. 提示：由于 $\{v_1, v_2, v_3\}$ 线性相关，你能写出一些方程且解出它.

33. 一种可能性是验证 T 不将零向量映射为零向量：$T(0,0)=(0,-3,0)$，但是，每个线性变换必须将零向量映射为零向量.

35. 取 $\mathbb{R}^3$ 中的 $\boldsymbol{u}$ 和 $\boldsymbol{v}$，c,d 为数，则 $c\boldsymbol{u}+d\boldsymbol{v}=(cu_1+dv_1, cu_2+dv_2, cu_3+dv_3)$，线性变换 T 是线性的，因为

$$\begin{aligned}T(c\boldsymbol{u}+d\boldsymbol{v}) &= (cu_1+dv_1, cu_2+dv_2, -(cu_3+dv_3)) \\ &= (cu_1+dv_1, cu_2+dv_2, -cu_3-dv_3) \\ &= (cu_1, cu_2, -cu_3)+(dv_1, dv_2, -dv_3) \\ &= c(u_1, u_2, -u_3)+d(v_1, v_2, -v_3) \\ &= cT(\boldsymbol{u})+dT(\boldsymbol{v})\end{aligned}$$

37. [M] 所有（-1,-1,1,0）的倍数.

39. [M] 是，对 $\boldsymbol{x}$ 的一个选择是（1,2,0,0）.

1.9

1. $\begin{bmatrix}3 & -5\\1 & 2\\3 & 0\\1 & 0\end{bmatrix}$ 3. $\begin{bmatrix}1 & 0\\-3 & 1\end{bmatrix}$ 5. $\begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}$

7. $\begin{bmatrix}-1/\sqrt{2} & 1/\sqrt{2}\\1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}$ 9. $\begin{bmatrix}0 & -1\\-1 & 0\end{bmatrix}$

11. 所描述的变换 T 将 $\boldsymbol{e}_1$ 映射为 $-\boldsymbol{e}_1$，将 $\boldsymbol{e}_2$ 映射为 $-\boldsymbol{e}_2$. 旋转 π 弧度也可以将 $\boldsymbol{e}_1$ 映射为 $-\boldsymbol{e}_1$，将 $\boldsymbol{e}_2$ 映射为 $-\boldsymbol{e}_2$. 因为一个线性变换完全由它对单位矩阵的列的映射而确定，这样的旋转变换对 $\mathbb{R}^2$ 中的每个向量的作用跟 T 是一样的.

13.

15. $\begin{bmatrix}2 & -4 & 0\\-1 & 0 & -1\\0 & -1 & 3\end{bmatrix}$ 17. $\begin{bmatrix}1 & 2 & 0 & 0\\0 & 0 & 0 & 0\\0 & 2 & 0 & 1\\0 & 1 & 0 & -1\end{bmatrix}$

19. $\begin{bmatrix}1 & -5 & 4\\0 & 1 & -6\end{bmatrix}$ 21. $\boldsymbol{x}=\begin{bmatrix}7\\-4\end{bmatrix}$

23. 查阅“学习指导”之前先回答问题.
对习题 25~28 验证你的答案.

25. 不是一对一的也不是将 $\mathbb{R}^4$ 映上到 $\mathbb{R}^4$ 的映射.

27. 不是一对一的，但是将 $\mathbb{R}^3$ 映上到 $\mathbb{R}^2$ 的映射

29. $\begin{bmatrix}\blacksquare & * & *\\0 & \blacksquare & *\\0 & 0 & \blacksquare\\0 & 0 & 0\end{bmatrix}$

31. n.（解释为什么，然后查阅“学习指导”.）

33. 提示：如果 $\boldsymbol{e}_j$ 是单位矩阵的第 j 列，那么 $B\boldsymbol{e}_j$ 是 B 的第 j 列.

35. 提示：$m>n$ 可能吗？对 $m<n$ 怎么样？

37. [M] 否（解释为什么）.

39. [M] 否（解释为什么）.

1.10

1. a. $x_1\begin{bmatrix}110\\4\\20\\2\end{bmatrix}+x_2\begin{bmatrix}130\\3\\18\\5\end{bmatrix}=\begin{bmatrix}295\\9\\48\\8\end{bmatrix}$，此处 x_1 是 Cheerios 份数，x_2 是 100%天然麦片的份数.

b. $\begin{bmatrix}110 & 130\\4 & 3\\20 & 18\\2 & 5\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}295\\9\\48\\8\end{bmatrix}$. 混合 1.5 份的 Cheerios 和 1 份 100%的天然麦片.

3. a. 0.99 苹果奶酪、1.54 西兰花和 0.79 鸡肉罐头
b. 1.09 全麦面和白切达奶酪，0.88 西兰花和 1.03 鸡肉罐头

5. $R\boldsymbol{i}=\boldsymbol{v}$, $\begin{bmatrix}11 & -5 & 0 & 0\\-5 & 10 & -1 & 0\\0 & -1 & 9 & -2\\0 & 0 & -2 & 10\end{bmatrix}\begin{bmatrix}I_1\\I_2\\I_3\\I_4\end{bmatrix}=\begin{bmatrix}50\\-40\\30\\-30\end{bmatrix}$

$$[\text{M}]: \boldsymbol{i}=\begin{bmatrix}I_1\\I_2\\I_3\\I_4\end{bmatrix}=\begin{bmatrix}3.68\\-1.90\\2.57\\-2.49\end{bmatrix}$$

7. $R\boldsymbol{i}=\boldsymbol{v}, \begin{bmatrix}12&-7&0&-4\\-7&15&-6&0\\0&-6&14&-5\\-4&0&-5&13\end{bmatrix}\begin{bmatrix}I_1\\I_2\\I_3\\I_4\end{bmatrix}=\begin{bmatrix}40\\30\\20\\-10\end{bmatrix}$

$$[\text{M}]: \boldsymbol{i}=\begin{bmatrix}I_1\\I_2\\I_3\\I_4\end{bmatrix}=\begin{bmatrix}11.43\\10.55\\8.04\\5.84\end{bmatrix}$$

9. $\boldsymbol{x}_{k+1}=M\boldsymbol{x}_k,\ k=0,1,2,\cdots$, 此处

$M=\begin{bmatrix}0.93&0.05\\0.07&0.95\end{bmatrix}$ 和 $\boldsymbol{x}_0=\begin{bmatrix}800\,000\\500\,000\end{bmatrix}$

2012 年的数量（$k=2$）是 $\boldsymbol{x}_2=\begin{bmatrix}741\,720\\558\,280\end{bmatrix}$

11. a. $M=\begin{bmatrix}0.983\,63&0.001\,67\\0.016\,37&0.998\,33\end{bmatrix}$

b. [M] $\boldsymbol{x}_6=\begin{bmatrix}30\,754\,500\\229\,449\,000\end{bmatrix}$

13. [M] a. 城市的人口减少，7 年后，人口会相等，但城市人口数量继续减少. 20 年后，城市人口仅有 417 000 人.（注意：417 456 在千位上四舍五入.）但每年的人口变化数量会越来越小.

b. 城市人口缓慢增长，郊区人口减少. 20 年后，城市人口将从 350 000 增长到 370 000.

第 1 章补充习题

1. a. F b. F c. T d. F e. T
f. T g. F h. F i. T j. F
k. T l. F m. T n. T o. T
p. T q. F r. T s. F t. T
u. F v. F w. F x. T y. T z. F

3. a. 任一具有 3 个方程的相容的方程组，其阶梯形为

$$\begin{bmatrix}\blacksquare&*&*&*\\0&\blacksquare&*&*\\0&0&0&0\end{bmatrix}\text{或}\begin{bmatrix}\blacksquare&*&*&*\\0&0&\blacksquare&*\\0&0&0&0\end{bmatrix}\text{或}\begin{bmatrix}0&\blacksquare&*&*\\0&0&\blacksquare&*\\0&0&0&0\end{bmatrix}$$

b. 任一具有 3 个方程的相容的方程组，其阶梯形为 I_3.

c. 任一具有 3 个变量和 3 个方程的不相容的线性方程组.

5. a. 解集：(i) 如果 $h=12$ 且 $k\neq 2$，是空集，(ii) 如果 $h\neq 12$，包含一个唯一解，(iii)如果 $h=12$ 且 $k=2$，则包含无穷多解.

b. 解集为空，如果 $k+3h=0$，否则，解集包含一个唯一解.

7. a. 令 $\boldsymbol{v}_1=\begin{bmatrix}2\\-5\\7\end{bmatrix}$, $\boldsymbol{v}_2=\begin{bmatrix}-4\\1\\-5\end{bmatrix}$, $\boldsymbol{v}_3=\begin{bmatrix}-2\\1\\-3\end{bmatrix}$ 和 $\boldsymbol{b}=\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}$.

“判断 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 是否可以生成 $\mathbb{R}^3$”，答案：否.

b. 令 $A=\begin{bmatrix}2&-4&-2\\-5&1&1\\7&-5&-3\end{bmatrix}$，“判断 A 的列是否可以生成 $\mathbb{R}^3$.”

c. 定义 $T(\boldsymbol{x})=A\boldsymbol{x}$，“判断 T 是否将 $\mathbb{R}^3$ 映上到 $\mathbb{R}^3$.”

9. $\begin{bmatrix}5\\6\end{bmatrix}=\frac{4}{3}\begin{bmatrix}2\\1\end{bmatrix}+\frac{7}{3}\begin{bmatrix}1\\2\end{bmatrix}$ 或 $\begin{bmatrix}5\\6\end{bmatrix}=\begin{bmatrix}8/3\\4/3\end{bmatrix}+\begin{bmatrix}7/3\\14/3\end{bmatrix}$.

10. 提示：在 x_1x_2 平面画出由 $\boldsymbol{a}_1$ 和 $\boldsymbol{a}_2$ 确定的格线.

11. 当方程组有一个自由变量时，解集是一条直线. 如果系数矩阵是 2×3，则三列中有两列是主元列. 例如，取系数矩阵 $\begin{bmatrix}1&2&*\\0&3&*\end{bmatrix}$，在第 3 列随便放值，所得矩阵是阶梯形. 对第 2 行做一次行变换以产生不是阶梯形的矩阵，例如 $\begin{bmatrix}1&2&1\\0&3&1\end{bmatrix}\sim\begin{bmatrix}1&2&1\\1&5&2\end{bmatrix}$.

12. 提示：方程 $A\boldsymbol{x}=\boldsymbol{0}$ 中有多少个自由变量？

13. $A=\begin{bmatrix}1&0&-3\\0&1&2\\0&0&0\end{bmatrix}$

15. a. 如果三个向量线性无关，则 a,c,f 必定全不为零.

b. 数 $a,\cdots,f$ 可取任意值.

16. 提示：从右向左将列写为 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_4$.

17. 提示：使用定理 7.

19. 设 M 为通过原点的直线且平行于 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 三点所连成的直线，则 $\boldsymbol{v}_2-\boldsymbol{v}_1$ 和 $\boldsymbol{v}_3-\boldsymbol{v}_1$ 都在 M 上，因此两个向量成倍数关系，比如 $\boldsymbol{v}_2-\boldsymbol{v}_1=k(\boldsymbol{v}_3-\boldsymbol{v}_1)$. 这个方程给出了线性相关关系：$(k-1)\boldsymbol{v}_1+\boldsymbol{v}_2-k\boldsymbol{v}_3=\boldsymbol{0}$. 第二种解法：直线的参数方程是 $\boldsymbol{x}=\boldsymbol{v}_1+t(\boldsymbol{v}_2-\boldsymbol{v}_1)$，由于 $\boldsymbol{v}_3$ 在直线上，存在一 t_0 使得 $\boldsymbol{v}_3=\boldsymbol{v}_1+t_0(\boldsymbol{v}_2-\boldsymbol{v}_1)=(1-t_0)\boldsymbol{v}_1+t_0\boldsymbol{v}_2$. 因此 $\boldsymbol{v}_3$ 是 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的线性组合，并且 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 线性相关.

21. $\begin{bmatrix}1&0&0\\0&-1&0\\0&0&1\end{bmatrix}$. 23. $a=4/5$, $b=-3/5$

25. a. 当建造 x_1 层的 A 设计时，向量列出有 3，2 和 1 个卧室的公寓数目.

b. $x_1\begin{bmatrix}3\\7\\8\end{bmatrix}+x_2\begin{bmatrix}4\\4\\8\end{bmatrix}+x_3\begin{bmatrix}5\\3\\9\end{bmatrix}$

c. [M] 利用 A 设计的 2 层和 B 设计的 15 层，或 A 设计的 6 层、B 设计的 2 层和 C 设计的 8 层. 这些是仅有的可行解. 也存在其他数值解，但它们需要一个或两个设计的负数的楼层，这没有实际意义.

第 2 章

2.1

1. $\begin{bmatrix}-4&0&2\\-8&10&-4\end{bmatrix}$, $\begin{bmatrix}3&-5&3\\-7&6&-7\end{bmatrix}$, 没有定义, $\begin{bmatrix}1&13\\-7&-6\end{bmatrix}$

3. $\begin{bmatrix}1&5\\-3&5\end{bmatrix}$, $\begin{bmatrix}6&-15\\9&-6\end{bmatrix}$

5. a. $A\boldsymbol{b}_1=\begin{bmatrix}-10\\0\\26\end{bmatrix}$, $A\boldsymbol{b}_2=\begin{bmatrix}11\\8\\-19\end{bmatrix}$, $AB=\begin{bmatrix}-10&11\\0&8\\26&-19\end{bmatrix}$

b. $AB=\begin{bmatrix}-1\cdot4+3(-2)&-1(-2)+3\cdot3\\2\cdot4+4(-2)&2(-2)+4\cdot3\\5\cdot4-3(-2)&2(-2)-3\cdot3\end{bmatrix}=\begin{bmatrix}-10&11\\0&8\\26&-19\end{bmatrix}$

7. 3×7

9. $k=-2$

11. $AD=\begin{bmatrix}5&6&6\\10&12&10\\15&15&12\end{bmatrix}$, $DA=\begin{bmatrix}5&10&15\\6&12&15\\6&10&12\end{bmatrix}$

被 D 右乘（即在右边乘）是用 D 相应的对角元素乘以 A 的每一列. 被 D 左乘是用 D 相应的对角元素乘以 A 的每一行. “学习指导”告诉你如何使 $AB=BA$，但你应该在查阅之前先自己试一试.

13. 提示：两个矩阵中的一个是 Q.

15. 在查阅“学习指导”之前回答这个问题.

17. $\boldsymbol{b}_1=\begin{bmatrix}3\\2\end{bmatrix}$, $\boldsymbol{b}_2=\begin{bmatrix}1\\4\end{bmatrix}$

19. AB 的第 3 列是 AB 前两列之和，以下是原因. 记 $B=[\boldsymbol{b}_1\quad\boldsymbol{b}_2\quad\boldsymbol{b}_3]$，根据定义，$AB$ 的第 3 列是 $A\boldsymbol{b}_3$，如果 $\boldsymbol{b}_3=\boldsymbol{b}_1+\boldsymbol{b}_2$，那么由矩阵向量乘法的性质可知 AB 的第三列是 $A\boldsymbol{b}_3=A(\boldsymbol{b}_1+\boldsymbol{b}_2)=A\boldsymbol{b}_1+A\boldsymbol{b}_2$.

21. A 的列是线性相关的，为什么？

23. 提示：假若 $\boldsymbol{x}$ 满足 $A\boldsymbol{x}=\boldsymbol{0}$，且证明 $\boldsymbol{x}$ 必须是 $\boldsymbol{0}$.

25. 提示：使用习题 23 和 24 的结果，并对乘积 CAD 使用乘法结合律.

27. $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}=\boldsymbol{v}\boldsymbol{u}^{\mathrm{T}}=-3a+2b-5c$

$\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}=\begin{bmatrix}-3a&-3b&-3c\\2a&2b&2c\\-5a&-5b&-5c\end{bmatrix}$, $\boldsymbol{v}\boldsymbol{u}^{\mathrm{T}}=\begin{bmatrix}-3a&2a&-5a\\-3b&2b&-5b\\-3c&2c&-5c\end{bmatrix}$

29. 提示：对定理 2（b），证明 $A(B+C)$ 中 (i,j) 元素和 $AB+AC$ 中的 (i,j) 元素相等.

31. 提示：利用乘积 I_mA 的定义和对 $\mathbb{R}^m$ 中的 $\boldsymbol{x}$ 有 $I_m\boldsymbol{x}=\boldsymbol{x}$ 的事实.

33. 提示：首先写出 $(AB)^{\mathrm{T}}$ 的 (i,j) 元素，它是 AB 的

(j,i) 元素，然后利用 B^{T} 中第 i 行的元素是 $b_{1i},\cdots,b_{ni}$，它们来自 B 的第 i 列，并且，A^{T} 中第 j 列的元素是 $a_{j1},\cdots,a_{jn}$，它们来自 A 的第 j 行，由行一列法则，$B^{\mathrm{T}}A^{\mathrm{T}}$ 中元素 (i,j) 是 $b_{1i}a_{j1}+\cdots+b_{ni}a_{jn}$.

35. [M]这里的答案依赖于你选取的矩阵程序，对 MATLAB，利用帮助命令查阅 zeros, ones, eye 和 diag. 对 TI-86 计算器，学习 dim,fill 和 iden 指令. 在 TI-86 中没有一个“diagonal”指令.
37. [M] 展示结果，并报告你的结论.
39. [M] 展示结果，并报告你的结论.
41. $\begin{bmatrix}1/3 & 1/3 & 1/3\\ 1/3 & 1/3 & 1/3\\ 1/3 & 1/3 & 1/3\end{bmatrix}$

2.2

1. $\begin{bmatrix}2 & -3\\ -5/2 & 4\end{bmatrix}$
3. $-\frac{1}{3}\begin{bmatrix}-3 & -3\\ 6 & 7\end{bmatrix}$ 或 $\begin{bmatrix}1 & 1\\ -2 & -7/3\end{bmatrix}$
5. $x_1=7, x_2=-9$
7. 对 $\boldsymbol{a}$ 和 $\boldsymbol{b}$，$\boldsymbol{x}$ 的解：$\begin{bmatrix}-9\\ 4\end{bmatrix},\begin{bmatrix}11\\ -5\end{bmatrix},\begin{bmatrix}6\\ -2\end{bmatrix},\begin{bmatrix}13\\ -5\end{bmatrix}$.
9. 在查阅“学习指导”之前写出你的答案.
11. 证明可以类似定理 5 的证明那样完成.
13. $AB=AC\Rightarrow A^{-1}AB=A^{-1}AC\Rightarrow IB=IC\Rightarrow B=C$. 否. 一般地，如果 A 不可逆，则 B 和 C 可能不同. 见 2.1 节习题 10.
15. 提示：对$[A,B]$应用初等矩阵的行最简形通常是单位矩阵 I.
17. $D=C^{-1}B^{-1}A^{-1}$，验证 D 满足要求.
19. 当找到 $X=CB-A$ 后，验证 X 是一个解.
21. 提示：考虑方程 $A\boldsymbol{x}=\boldsymbol{0}$.
23. 提示：如果 $A\boldsymbol{x}=\boldsymbol{0}$ 仅有平凡解，那么方程 $A\boldsymbol{x}=\boldsymbol{0}$ 没有自由变量，且 A 的每个列是一主元列.
25. 提示：考虑当 $a=b=0$ 的情形，然后考虑向量 $\begin{bmatrix}-b\\ a\end{bmatrix}$ 且利用事实 $ad-bc=0$.
27. 提示：对（a），将 2.1 节中例 6 后的方框中公式中的 A 和 B 互换位置，然后用单位矩阵代替 B. 对（b）和（c），先将 A 写为 $A=\begin{bmatrix}\mathrm{row}_1(A)\\ \mathrm{row}_2(A)\\ \mathrm{row}_3(A)\end{bmatrix}$.
29. $\frac{1}{3}\begin{bmatrix}-9 & 3\\ -4 & 1\end{bmatrix}$
31. $\begin{bmatrix}8 & 3 & 1\\ 10 & 4 & 1\\ 7/2 & 3/2 & 1/2\end{bmatrix}$
33. $A^{-1}=B=\begin{bmatrix}1 & 0 & 0 & \cdots & 0\\ -1 & 1 & 0 & & 0\\ 0 & -1 & 1 & & \\ \vdots & & & \ddots & \vdots\\ 0 & 0 & \cdots & -1 & 1\end{bmatrix}$. 提示：对 $j=1,\cdots,n$，设 $\boldsymbol{a}_j,\boldsymbol{b}_j,\boldsymbol{e}_j$ 分别表示 A,B,I 的第 j 列，利用对 $j=1,\cdots,n-1$，有 $\boldsymbol{a}_j-\boldsymbol{a}_{j+1}=\boldsymbol{e}_j$ 和 $\boldsymbol{b}_j=\boldsymbol{e}_j-\boldsymbol{e}_{j+1}$，并且 $\boldsymbol{a}_n=\boldsymbol{b}_n=\boldsymbol{e}_n$.
35. $\begin{bmatrix}3\\ 0\\ -1\end{bmatrix}$，通过行化简$[A\quad \boldsymbol{e}_3]$求得.
37. $C=\begin{bmatrix}1 & 1 & -1\\ -1 & 1 & 0\end{bmatrix}$
39. 分别是 0.62，0.66 和 0.52 英寸.
41. [M]分别为 0.95，6.19，11.43 和 3.81 牛顿的力.

2.3

为节约答案的篇幅，我们用 IMT 表示可逆矩阵定理（定理 8）.

1. 可逆，由 IMT，列线性无关，因为矩阵的任一列不是其他列的倍数，并且，由 2.2 节定理 4，行列式不等于零，因此矩阵是可逆的.
3. 注意到 A^T 的每一列有一个主元，所以由 IMT 知 A^T可逆，因此由 IMT 知，A 也是可逆的.
5. 由 IMT，因为矩阵有一列为零，所以列向量之间线性相关，因此矩阵是不可逆矩阵.
7. 可逆，由 IMT，矩阵是行变换后变为

$$\begin{bmatrix} -1 & -3 & 0 & 1 \\ 0 & -4 & 8 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$ 且有 4 个主元位置.

9. [M] 4×4 矩阵有 4 个主元位置，由 IMT，矩阵可逆.

11. “学习指导”会帮助你，但首先要建立在仔细阅读课本的基础上，尝试回答问题.

13. 一个上三角形方阵是可逆的充分必要条件是所有在对角线上的元素非零. 为什么？

注意：习题 15~29 的答案与 IMT 有关. 在许多时候，部分或全部的答案可以是基于用来建立 IMT 的结论.

15. 不可逆，因为 IMT 的命题(h)是假的. 一个 4×4 矩阵的列不能生成 $\mathbb{R}^4$ 时，这个矩阵不可能是可逆的.

17. 如果 A 有两个相同的列，则它的列是线性相关的. IMT 的命题（e）说明 A 不可能是可逆的.

19. 根据 IMT 中的命题（e），D 是可逆的. 于是由 IMT 中的命题（g），方程 $D\boldsymbol{x}=\boldsymbol{b}$ 对所有属于 $\mathbb{R}^7$ 的 $\boldsymbol{b}$ 有一个解，你有更多结论吗？

21. 由 2.2 节定理 5 或 IMT 下面的段落，矩阵 C 不可能是可逆的. 因此 IMT 中的命题（g）是假的，同样（h）也是假的，即 C 的列不能生成 $\mathbb{R}^n$.

23. 对 F，IMT 的命题（g）是假的，因此命题（d）也是假的，即 $F\boldsymbol{x}=\boldsymbol{0}$ 有一个非平凡解.

25. 提示：首先利用 IMT.

27. 由于 AB 是可逆的，存在一个矩阵 W 使得 $ABW=I$ 且有 $A(BW)=I$，但是，由这个计算本身不能证明 A 是可逆，为什么不能？在查阅“学习指导”之前完成证明.

29. 因为变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 是一对一的，IMT 的命题（f）是真的，因此命题（i）也是真的，变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 将 $\mathbb{R}^n$ 映上到 $\mathbb{R}^n$，并且 A 是可逆的，由此根据定理 9 推出变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 是可逆的.

31. 提示：如果对任一 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 有一个解，那么 A 的每一行有一个主元（根据 1.4 节定理 4）. 方程 $A\boldsymbol{x}=\boldsymbol{b}$ 是否有自由变量？

33. 提示：首先证明 T 的标准矩阵是可逆的，然后使用定理证明 $T^{-1}(\boldsymbol{x})=B\boldsymbol{x}$，其中 $B=\begin{bmatrix} 7 & 9 \\ 4 & 5 \end{bmatrix}$.

35. 提示：为了证明 T 是一对一的，假设对 $\mathbb{R}^n$ 中的向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$，有 $T(\boldsymbol{u})=T(\boldsymbol{v})$，证明 $\boldsymbol{u}=\boldsymbol{v}$. 为了证明 T 是映上的，假设 $\boldsymbol{y}$ 表示 $\mathbb{R}^n$ 中的任一向量，利用 S 的逆来产生一个向量 $\boldsymbol{x}$ 使得 $T(\boldsymbol{x})=\boldsymbol{y}$. 第二种证明可以使用定理 9 及 1.9 节中的一个定理.

37. 提示：考虑 T 和 U 的标准矩阵.

39. 如果 T 将 $\mathbb{R}^n$ 映射到 $\mathbb{R}^n$，那么根据 1.9 节的定理 12，其标准矩阵 A 的列生成 $\mathbb{R}^n$. 由 IMT 知，A 是可逆的. 因此根据定理 9，T 是可逆的，并且 A^{-1} 是 T^{-1} 的标准矩阵. 因为 A^{-1} 也可逆，由 IMT 知，它的列是线性无关的，并且生成 $\mathbb{R}^n$. 对变换 T^{-1} 应用 1.9 节的定理 12 证明 T^{-1} 是从 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 的一对一映射.

41. [M] a.（3）的精确解是 $x_1=3.94$ 和 $x_2=0.49$. （4）的精确解是 $x_1=2.90$ 和 $x_2=2.00$.

    b. 当用（4）的解作为（3）的一个近似解，近似解 x_1 取 2.90 时误差为 26%，近似解 x_2 取 2.0 时误差为 308%.

    c. 系数矩阵的条件是 3363，从（3）和（4）的解改变的百分比大约是 7700 倍方程右端改变的百分比. 这与条件数是同样阶的大小. 条件数大约给出方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的解随 $\boldsymbol{b}$ 改变的敏感性的粗略度量. 更进一步关于条件数的信息在第 6 章的最后和第 7 章中给出.

43. [M] $\text{cond}(A)\approx 69\,000$，位于 10^4 和 10^5 之间，所以大约 4 或 5 位数字的精度可能丢失，一些用 MATLAB 的实验可以验证，$\boldsymbol{x}$ 和 $\boldsymbol{x}_1$ 有大约 11 或 12 位数字相一致.

45. [M]当要求对一个阶数为 12 或更大的使用浮点数的希尔伯特矩阵求逆时，MATLAB 给出一个警告，乘积 AA^{-1} 的非对角元素会有非零值. 如果没有出现这种情况，尝试一个更大的矩阵.

2.4

1. $\begin{bmatrix} A & B \\ EA+C & EB+D \end{bmatrix}$ 3. $\begin{bmatrix} C & D \\ A & B \end{bmatrix}$

5. $Y = B^{-1}$（解释为什么），$X = -B^{-1}A, Z = C$

7. $X = A^{-1}$（为什么？），$Y = -BA^{-1}, Z = 0$（为什么？）

9. $A_{21} = -B_{21}B_{11}^{-1}$，$A_{31} = -B_{31}B_{11}^{-1}$，

$C_{22} = B_{22} - B_{21}B_{11}^{-1}B_{12}$

11. 你可查阅“学习指导”中的参考.

13. 提示：若 A 是可逆的，且取 $A^{-1} = \begin{bmatrix} D & E \\ F & G \end{bmatrix}$，证明 $BD = I$ 和 $CG = I$，这说明 B 和 C 可逆.（解释为什么！）相反，若 B 和 C 可逆，为证明 A 可逆，猜测 A^{-1} 是什么并检查.

15. $G_{k+1} = [X_k \ \ \boldsymbol{x}_{k+1}] \begin{bmatrix} X_k^{\mathrm{T}} \\ \boldsymbol{x}_{k+1}^{\mathrm{T}} \end{bmatrix} = X_k X_k^{\mathrm{T}} + \boldsymbol{x}_{k+1}\boldsymbol{x}_{k+1}^{\mathrm{T}}$

$= G_k + \boldsymbol{x}_{k+1}\boldsymbol{x}_{k+1}^{\mathrm{T}}$

只有外积矩阵 $\boldsymbol{x}_{k+1}\boldsymbol{x}_{k+1}^{\mathrm{T}}$ 需要计算（然后加到 G_k 上）.

17. $\begin{bmatrix} I & 0 \\ X & I \end{bmatrix}$的逆为$\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}$，类似地，$\begin{bmatrix} I & Y \\ 0 & I \end{bmatrix}$存在逆. 由方程（7），可得

$$\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} I & -Y \\ 0 & I \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ 0 & S \end{bmatrix} \quad (*)$$

若 A 是可逆的，则在（*）式右边的矩阵是可逆矩阵的积，因此可逆. 由习题 13，A_{11} 与 S 一定是可逆的.

19. $W(s) = I_m - C(A - sI_n)^{-1}B$. 这是系统矩阵 $A - sI_n$ 的舒尔补.

21. a. $A^2 = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 1+0 & 0+0 \\ 2-2 & 0+(-1)^2 \end{bmatrix}$

$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

b. $M^2 = \begin{bmatrix} A & 0 \\ I & -A \end{bmatrix}\begin{bmatrix} A & 0 \\ I & -A \end{bmatrix}$

$= \begin{bmatrix} A^2+0 & 0+0 \\ A-A & 0+(-A)^2 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

23. 如果 A_1 和 B_1 是 $(k+1)\times(k+1)$ 且为下三角形，那么我们可以写出 $A_1 = \begin{bmatrix} a & \boldsymbol{0}^{\mathrm{T}} \\ \boldsymbol{v} & A \end{bmatrix}$ 和 $B_1 = \begin{bmatrix} b & \boldsymbol{0}^{\mathrm{T}} \\ \boldsymbol{w} & B \end{bmatrix}$，此处 A 和 B 是 $k\times k$ 且为下三角形，$\boldsymbol{v}$ 和 $\boldsymbol{w}$ 属于 $\mathbb{R}^k$，且 a 和 b 是适合的数，假若任意 $k\times k$ 下三角形矩阵的乘积是下三角形矩阵，然后计算乘积 A_1B_1，你能得出什么结论？

25. 利用习题 13 求形如 $B = \begin{bmatrix} B_{11} & 0 \\ 0 & B_{22} \end{bmatrix}$ 的矩阵的逆，其中 B_{11} 是 $p\times p$，B_{22} 是 $q\times q$，B 可逆. 将矩阵 A 分块，用两遍你的结果求得

$$A^{-1} = \begin{bmatrix} -5 & 2 & 0 & 0 & 0 \\ 3 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 3 & -4 \\ 0 & 0 & 0 & -5/2 & 7/2 \end{bmatrix}.$$

27. [M]a, b 习题中要使用的命令依赖于所使用的矩阵程序.

c. 从分块矩阵方程$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix}$，其中 $\boldsymbol{x}_1, \boldsymbol{b}_1$ 属于 $\mathbb{R}^{20}$，$\boldsymbol{x}_2, \boldsymbol{b}_2$ 属于 $\mathbb{R}^{30}$，可以得到 $A_{11}\boldsymbol{x}_1 = \boldsymbol{b}_1$，从而可以求出 $\boldsymbol{x}_1$，从方程 $A_{21}\boldsymbol{x}_1 + A_{22}\boldsymbol{x}_2 = \boldsymbol{b}_2$ 得到 $A_{22}\boldsymbol{x}_2 = \boldsymbol{b}_2 - A_{21}\boldsymbol{x}_1$，再通过行简化矩阵 $[A_{22} \ \ \boldsymbol{c}]$ 求出 $\boldsymbol{x}_2$，其中 $\boldsymbol{c} = \boldsymbol{b}_2 - A_{21}\boldsymbol{x}_1$.

2.5

1. $L\boldsymbol{y} = \boldsymbol{b} \Rightarrow \boldsymbol{y} = \begin{bmatrix} -7 \\ -2 \\ 6 \end{bmatrix}$, $U\boldsymbol{x} = \boldsymbol{y} \Rightarrow \boldsymbol{x} = \begin{bmatrix} 3 \\ 4 \\ -6 \end{bmatrix}$

3. $\boldsymbol{y} = \begin{bmatrix} 6 \\ 12 \\ 0 \end{bmatrix}$, $\boldsymbol{x} = \begin{bmatrix} -5 \\ -4 \\ 0 \end{bmatrix}$ 5. $\boldsymbol{y} = \begin{bmatrix} 1 \\ 3 \\ 1 \\ -4 \end{bmatrix}$, $\boldsymbol{x} = \begin{bmatrix} 38 \\ 16 \\ 8.5 \\ -4 \end{bmatrix}$

7. $LU=\begin{bmatrix}1&0\\-3/2&1\end{bmatrix}\begin{bmatrix}2&5\\0&7/2\end{bmatrix}$

9. $\begin{bmatrix}1&0&0\\-3&1&0\\3&2&1\end{bmatrix}\begin{bmatrix}3&1&2\\0&3&2\\0&0&4\end{bmatrix}$

11. $\begin{bmatrix}1&0&0\\2&1&0\\-1&1&1\end{bmatrix}\begin{bmatrix}3&7&2\\0&5&0\\0&0&5\end{bmatrix}$

13. $\begin{bmatrix}1&0&0&0\\-1&1&0&0\\4&5&1&0\\-2&-1&0&1\end{bmatrix}\begin{bmatrix}1&3&-5&-3\\0&-2&3&1\\0&0&0&0\\0&0&0&0\end{bmatrix}$

15. $\begin{bmatrix}1&0&0\\-3&1&0\\2&3&1\end{bmatrix}\begin{bmatrix}2&0&5&2\\0&3&2&3\\0&0&0&4\end{bmatrix}$

17. $U^{-1}=\begin{bmatrix}1/2&-3/4&-2\\0&-1/4&-1\\0&0&-1/2\end{bmatrix}$,

$L^{-1}=\begin{bmatrix}1&0&0\\2&1&0\\-2&-1&1\end{bmatrix}$,

$A^{-1}=\begin{bmatrix}3&5/4&-2\\3/2&3/4&-1\\1&1/2&-1/2\end{bmatrix}$

19. 提示：考虑行简化$[A\ \ I]$.

21. 提示：将行变换表示为一系列初等矩阵的乘积.

23. a. 将 D 的行表示为列向量的转置. 那么从分块矩阵的乘法得到：

$$A=CD=[\boldsymbol{c}_1\ \cdots\ \boldsymbol{c}_4]\begin{bmatrix}\boldsymbol{v}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{v}_4^{\mathrm{T}}\end{bmatrix}=\boldsymbol{c}_1\boldsymbol{v}_1^{\mathrm{T}}+\cdots+\boldsymbol{c}_4\boldsymbol{v}_4^{\mathrm{T}}$$

b. A 有 40 000 个元素，由于 C 有 1600 个元素且 D 有 400 个元素，两者合在一起只占存储 A 所需的 5%.

25. 解释为什么 U,D 和 V^{T} 是可逆的，那么对可逆矩阵的乘积的逆运用一个定理.

27. a.

i_1 i_2 i_2 i_3
v_1 1/2 Ω v_2 9/2 Ω v_3

b.

i_1 i_2 i_2 i_3
v_1 6 Ω v_2 3/4 Ω v_3

29. a. $\begin{bmatrix}1+R_3/R_2 & -R_1-R_3-(R_1R_3)/R_2\\ -1/R_2 & 1+R_1/R_2\end{bmatrix}$

b. $A=\begin{bmatrix}3&-12\\-1/3&5/3\end{bmatrix}=\begin{bmatrix}1&-6\\0&1\end{bmatrix}\begin{bmatrix}1&0\\-1/3&1\end{bmatrix}\begin{bmatrix}1&-2\\0&1\end{bmatrix}$

设 $R_1=2\Omega$，$R_2=3\Omega$，$R_3=6\Omega$.

31. [M]

a.

$$L=\begin{bmatrix}1&0&0&0&0&0&0&0\\-0.25&1&0&0&0&0&0&0\\-0.25&-0.0667&1&0&0&0&0&0\\0&-0.2667&-0.2857&1&0&0&0&0\\0&0&-0.2679&-0.0833&1&0&0&0\\0&0&0&-0.2917&-0.2921&1&0&0\\0&0&0&0&-0.2697&-0.0861&1&0\\0&0&0&0&0&-0.2948&-0.2931&1\end{bmatrix}$$

$$U=\begin{bmatrix}4&-1&-1&0&0&0&0&0\\0&3.75&-0.25&-1&0&0&0&0\\0&0&3.7333&-1.0667&-1&0&0&0\\0&0&0&3.4286&-0.2857&-1&0&0\\0&0&0&0&3.7083&-1.0833&-1&0\\0&0&0&0&0&3.3919&-0.2921&-1\\0&0&0&0&0&0&3.7052&-1.0861\\0&0&0&0&0&0&0&3.3868\end{bmatrix}$$

b. $x=(3.9569,\ 6.5885,\ 4.2392,\ 7.3971,\ 5.6029,\ 8.7608,\ 9.4115,\ 12.0431)$

c.

$$A^{-1}=\begin{bmatrix}0.2953&0.0866&0.0945&0.0509&0.0318&0.0227&0.0100&0.0082\\0.0866&0.2953&0.0509&0.0945&0.0227&0.0318&0.0082&0.0100\\0.0945&0.0509&0.3271&0.1093&0.1045&0.0591&0.0318&0.0227\\0.0509&0.0945&0.1093&0.3271&0.0591&0.1045&0.0227&0.0318\\0.0318&0.0227&0.1045&0.0591&0.3271&0.1093&0.0945&0.0509\\0.0227&0.0318&0.0591&0.1045&0.1093&0.3271&0.0509&0.0945\\0.0100&0.0082&0.0318&0.0227&0.0945&0.0509&0.2953&0.0866\\0.0082&0.0100&0.0227&0.0318&0.0509&0.0945&0.0866&0.2953\end{bmatrix}$$

可直接得到 A^{-1} 然后计算 $A^{-1}-U^{-1}L^{-1}$，比较两种计算矩阵逆的方法.

2.6

1. $C=\begin{bmatrix}0.10&0.60&0.60\\0.30&0.20&0\\0.30&0.10&0.10\end{bmatrix}$, $\{\text{中间需求}\}=\begin{bmatrix}60\\20\\10\end{bmatrix}$

3. $x=\begin{bmatrix}44.44\\16.67\\16.67\end{bmatrix}$　　5. $x=\begin{bmatrix}110\\120\end{bmatrix}$

7. a. $\begin{bmatrix}1.6\\1.2\end{bmatrix}$　　b. $\begin{bmatrix}111.6\\121.2\end{bmatrix}$

c. 提示：求$(I-C)^{-1}$的公式. 见“学习指导”

9. $x=\begin{bmatrix}82.8\\131.0\\110.3\end{bmatrix}$

11. 提示：利用转置的性质得到 $p^T=p^TC+v^T$，使得 $p^Tx=(p^TC+v^T)x=p^TCx+v^Tx$ 现在用乘积方程计算 p^Tx.

13. [M] $x=(99\,576, 97\,703, 51\,231, 131\,570, 49\,488, 329\,554, 13\,835)$. 答案 x 中的元素比在 d 中只精确到接近千位的元素更准确，所以，更实际的 x 的答案是 $x=1000\times(100, 98, 51, 132, 49, 330, 14)$.

15. [M] $x^{(12)}$ 是第一个向量，其元素精确到接近千位. $x^{(12)}$ 的计算需要 1260 次浮算，而 $[(I-C)\ \ d]$ 的行变换仅花费 550 次浮算. 如果 C 是大于 20×20，那么用迭代计算 $x^{(12)}$ 比用行变换计算平衡向量 x 需要更少的浮算. 随着 C 的阶数增加，迭代方法的优势会更加明显. 而且，在经济学的较大模型中的 C 变得更稀疏，较少的迭代就能得到所需要的精度.

2.7

1. $\begin{bmatrix}1&0.25&0\\0&1&0\\0&0&1\end{bmatrix}$　　3. $\begin{bmatrix}0&-1&-1\\1&0&2\\0&0&1\end{bmatrix}$

5. $\begin{bmatrix}1/\sqrt{2}&1/\sqrt{2}&0\\1/\sqrt{2}&-1/\sqrt{2}&0\\0&0&1\end{bmatrix}$

7. $\begin{bmatrix}1/2&-\sqrt{3}/2&3+4\sqrt{3}\\\sqrt{3}/2&1/2&4-3\sqrt{3}\\0&0&1\end{bmatrix}$

见练习题.

9. $A(BD)$ 需要 800 次乘法，$(AB)D$ 需要 408 次乘法，第一种方法乘法次数大约是第二种方法的 2 倍. 如果 D 有 10 000 列，两种计算次数分别为 80 000 和 40 008 次.

11. 利用以下事实：

$$\sec\varphi-\tan\varphi\sin\varphi=\frac{1}{\cos\varphi}-\frac{\sin^2\varphi}{\cos\varphi}=\cos\varphi$$

13. $\begin{bmatrix}A&p\\0^T&1\end{bmatrix}=\begin{bmatrix}I&p\\0^T&1\end{bmatrix}\begin{bmatrix}A&0\\0^T&1\end{bmatrix}$，首先施加线性变换 A，然后对之平移 p.

15. (12,−6,−3)

17. $\begin{bmatrix}1&0&0&0\\0&1/2&-\sqrt{3}/2&0\\0&\sqrt{3}/2&1/2&0\\0&0&0&1\end{bmatrix}$

19. 三角形顶点在 (7,2,0)，(7.5,5,0)，(5,5,0).

21. [M] $\begin{bmatrix}2.258\,6&-1.039\,5&-0.347\,3\\-1.349\,5&2.344\,1&0.069\,6\\0.091\,0&-0.304\,6&1.277\,7\end{bmatrix}\begin{bmatrix}X\\Y\\Z\end{bmatrix}=\begin{bmatrix}R\\G\\B\end{bmatrix}$、

2.8

1. 集合对加法封闭，但对一个负数，标量乘法不封闭. （给出一个简单例子.）

3. 集合在加法或标量乘法下不封闭.

5. 是. 对应于 $[v_1\ \ v_2\ \ w]$ 的方程组相容.

7. a. 3 个向量 v_1, v_2, v_3.

b. 无穷多向量.

c. 是，因为方程 $Ax=p$ 有一个解.

9. 否，因为 $Ap\neq 0$.

11. $p=4,\ q=3$, Nul$A$ 是 $\mathbb{R}^4$ 的一个子空间，因为 $Ax=0$ 的解必然有四个元素，以便与的列相匹配. Col$A$ 是 $\mathbb{R}^3$ 的一个子空间，因为每个列向量有三个元素.

13. 对 Nul A 作为例子选 (1,−2,1,0) 或 (−1,4,0,1)，

对 Col A，选 A 的任意一列.

15. 设 A 是用所给向量做为列的矩阵，因为行列式不等于零，因此 A 是可逆的，根据 IMT（或例 5），其列构成 $\mathbb{R}^2$ 的一组基.（也可给出 A 可逆的其他理由.）

17. 是，设 A 是列为给定向量的矩阵，A 的行变换表明有 3 个主元. 所以 A 是可逆的，根据 IMT，A 的列构成 $\mathbb{R}^3$ 的一个基.

19. 设 A 是 3×2 矩阵，它的列是给定的向量，A 的列不能张成 $\mathbb{R}^3$，因为 A 中不是每一行有一个主元，所以列不能构成 $\mathbb{R}^3$ 的一个基.（它们是 $\mathbb{R}^3$ 中的平面的一个基.）

21. 仔细阅读这一节，然后查阅“学习指导”之前写出你的答案，这一节内容和关键概念你必须现在学习，然后才能继续学习下面的内容.

23. Col A 的基：$\begin{bmatrix}4\\6\\3\end{bmatrix},\begin{bmatrix}5\\5\\4\end{bmatrix}$；Nul A 的基：$\begin{bmatrix}4\\-5\\1\\0\end{bmatrix},\begin{bmatrix}-7\\6\\0\\1\end{bmatrix}$

25. Col A 的基：$\begin{bmatrix}1\\-1\\-2\\3\end{bmatrix},\begin{bmatrix}4\\2\\2\\6\end{bmatrix},\begin{bmatrix}-3\\3\\5\\-5\end{bmatrix}$；

Nul A 的基：$\begin{bmatrix}2\\-2.5\\1\\0\\0\end{bmatrix},\begin{bmatrix}-7\\0.5\\0\\-4\\1\end{bmatrix}$

27. 构造一个非零的 3×3 矩阵 A，用 A 的列的任意方便的线性组合作为 $\boldsymbol{b}$.

29. 提示：你需要一个非零的矩阵，其列是线性相关的.

31. 如果 Col $F \neq \mathbb{R}^5$，则 F 的列不能生成 $\mathbb{R}^5$. 由于 F 是方阵，IMT 说明 F 是不可逆的，方程 $F\boldsymbol{x}=\boldsymbol{0}$ 有非平凡解，即 Nul F 包含非零向量. 另外的描述方式是 Nul $F \neq \{\boldsymbol{0}\}$.

33. 如果 Nul $C=\{\boldsymbol{0}\}$，则方程 $C\boldsymbol{x}=\boldsymbol{0}$ 只有平凡解. 由于 C 是方阵，IMT 说明 C 是可逆的，方程 $C\boldsymbol{x}=\boldsymbol{b}$ 对 $\mathbb{R}^6$ 中的每个 $\boldsymbol{b}$ 都有一个解，而且由 2.2 节定理 5，每个解是唯一的.

35. 如果 Nul B 包含非零向量，则方程 $B\boldsymbol{x}=\boldsymbol{0}$ 有非平凡解，因为 B 是方阵，IMT 表明 B 不可逆并且 B 的列不能生成 $\mathbb{R}^5$，所以 Col B 是 $\mathbb{R}^5$ 的一个子空间，但是 Col $B \neq \mathbb{R}^5$.

37. [M]写出 A 的简化阶梯形，将 A 的主元列选为 Col A 的一个基. 对于 Nul A，写出 $A\boldsymbol{x}=\boldsymbol{0}$ 的解的参数向量形式.

Col A 的基：$\begin{bmatrix}3\\-7\\-5\\3\end{bmatrix},\begin{bmatrix}-5\\9\\7\\-7\end{bmatrix}$；

Nul A 的基：$\begin{bmatrix}-2.5\\-1.5\\1\\0\\0\end{bmatrix},\begin{bmatrix}4.5\\2.5\\0\\1\\0\end{bmatrix},\begin{bmatrix}-3.5\\-1.5\\0\\0\\1\end{bmatrix}$

2.9

1. $\boldsymbol{x}=3\boldsymbol{b}_1+2\boldsymbol{b}_2=3\begin{bmatrix}1\\1\end{bmatrix}+2\begin{bmatrix}2\\-1\end{bmatrix}=\begin{bmatrix}7\\1\end{bmatrix}$

3. $\begin{bmatrix}1\\2\end{bmatrix}$

5. $\begin{bmatrix}4\\1\end{bmatrix}$

7. $[\boldsymbol{w}]_B=\begin{bmatrix}2\\-1\end{bmatrix}$，$[\boldsymbol{x}]_B=\begin{bmatrix}1.5\\0.5\end{bmatrix}$

9. Col A 的基：$\begin{bmatrix}1\\3\\2\\5\end{bmatrix},\begin{bmatrix}2\\1\\-1\\0\end{bmatrix},\begin{bmatrix}-6\\5\\9\\14\end{bmatrix}$；$\dim \text{Col } A=3$

Nul A 的基：$\begin{bmatrix}-3\\1\\0\\0\end{bmatrix}$；$\dim \text{Nul } A=1$

11. Col A 的基：$\begin{bmatrix} 2 \\ 3 \\ 0 \\ -3 \end{bmatrix}, \begin{bmatrix} -5 \\ -8 \\ 9 \\ -7 \end{bmatrix}$;dim Col $A = 2$

Nul A 的基：$\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$;dim Nul $A = 3$

13. 子空间 H 的基为向量 $\boldsymbol{v}_1$，$\boldsymbol{v}_2$，$\boldsymbol{v}_3$，因此 $\dim H = 3$.

15. Col $A = \mathbb{R}^4$，因为 A 的每一行有一个主元，A 的列生成 $\mathbb{R}^4$. Nul A 不等于 $\mathbb{R}^2$，因为 Nul A 是 $\mathbb{R}^6$ 的子空间. 但 Nul A 是二维的，理由：方程 $A\boldsymbol{x} = \boldsymbol{0}$ 有两个自由变量，因为 A 有 6 列，当中只有 4 列是主元列.

17. 在给出你的理由后再查看“学习指导”.

19. $A\boldsymbol{x} = \boldsymbol{0}$ 的解空间有一组三个向量组成的基，意味着 dim Nul $A = 3$. 因为 5×7 矩阵 A 有 7 列，秩定理说明 rank $A = 7 -$ dim Nul $A = 4$. 见“学习指导”中不用秩定理来论证.

21. 一个 9×8 矩阵有 8 列，根据秩定理，dim Nul $A = 8 -$ rank A. 因为其秩是 7，故 dim Nul $A = 1$，即 $A\boldsymbol{x} = \boldsymbol{0}$ 的解空间的维数是 1.

23. 创建一个 3×5 矩阵 A，且具有 2 个主元列. 剩下的三列对应于方程 $A\boldsymbol{x}=\boldsymbol{0}$ 中的自由变量.

25. 由定义，A 的 p 列生成 Col A. 如果 dim Col $A = p$，则根据基定理，p 列的生成集自然成为 Col A 的一组基，特别地，这些列是线性无关的.

27. a. 提示：B 的列生成 W，每个向量 $\mathbf{a}_j$ 属于 W. 向量 $\boldsymbol{c}_j$ 属于 $\mathbb{R}^p$ 是因为 B 有 p 列.

b. 提示：C 的维数是多少？

c. 提示：B 和 C 与 A 关系如何？

29. [M]你的计算应该说明 $[\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \boldsymbol{x}]$ 对应于一个相容的方程组，$\boldsymbol{x}$ 的 $\mathcal{B}-$坐标向量是（2,–1）.

第 2 章补充习题

1. a. T　b. F　c. T　d. F
e. F　f. F　g. T　h. T
i. T　j. F　k. T　l. F
m. F　n. T　o. F　p. T

3. I

5. $A^2 = 2A - I$，乘以 A 得到：$A^3 = 2A^2 - A$.
将 $A^2 = 2A - I$ 代入：$A^3 = 2(2A - I) - A = 3A - 2I$.
再乘以 A：$A^4 = A(3A - 2I) = 3A^2 - 2A$.
再将 $A^2 = 2A - I$ 代入一次：
$A^4 = 3(2A - I) - 2A = 4A - 3I$.

7. $\begin{bmatrix} 10 & -1 \\ 9 & 10 \\ -5 & -3 \end{bmatrix}$.

9. $\begin{bmatrix} -3 & 13 \\ -8 & 27 \end{bmatrix}$

11. a. $p(x_i) = c_0 + c_1 x_i + \cdots + c_{n-1} x_i^{n-1}$

$$= \text{row}_i(V) \cdot \begin{bmatrix} c_0 \\ \vdots \\ c_{n-1} \end{bmatrix} = \text{row}_i(V\boldsymbol{c}) = y_i$$

b. 假若 $x_1, \cdots, x_n$ 是不同的，并设对某向量 $\boldsymbol{c}$ 有 $V\boldsymbol{c} = \boldsymbol{0}$，那么 $\boldsymbol{c}$ 中的元素是一个在 $x_1, \cdots, x_n$ 处为零的多项式的系数，但是，一个 $n-1$ 次多项式不可能有 n 个零点，因此这个多项式必然恒等于零，也就是 $\boldsymbol{c}$ 中的元素必须都是零，这表明 V 中的列是线性无关的.

c. 提示：当 $x_1, \cdots, x_n$ 是不同的，存在一个向量 $\boldsymbol{c}$ 使得 $V\boldsymbol{c} = \boldsymbol{y}$. 为什么？

13. a. $P^2 = (\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}})(\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}) = \boldsymbol{u}(\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u})\boldsymbol{u}^{\mathrm{T}} = \boldsymbol{u}(1)\boldsymbol{u}^{\mathrm{T}} = P$

b. $P^{\mathrm{T}} = (\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{u}^{\mathrm{TT}}\boldsymbol{u}^{\mathrm{T}} = \boldsymbol{u}\boldsymbol{u}^{\mathrm{T}} = P$

c. $Q^2 = (I - 2P)(I - 2P)$　　因为（a）
$= I - I(2P) - 2PI + 2P(2P)$
$= I - 4P + 4P^2 = I$

15. 用初等矩阵左乘产生一个行初等变换：
$B \sim E_1B \sim E_2E_1B \sim E_3E_2E_1B = C$，所以 B 是行等

价于 C，因为行变换是可逆的，C 行等价于 B（换言之，利用 E_i 的逆作行变换可证明 C 被变为 B．）

17. 由于 B 是 4×6（列数多于行数），它的 6 列线性相关，且有非零 $\mathbf{x}$ 使得 $B\mathbf{x}=\mathbf{0}$，于是 $AB\mathbf{x}=A\mathbf{0}=\mathbf{0}$，由可逆矩阵定理，说明矩阵 AB 是不可逆的．

19. [M]对 4 位小数，当 k 增加时，

$$A^k \to \begin{bmatrix} 0.2857 & 0.2857 & 0.2857 \\ 0.4286 & 0.4286 & 0.4286 \\ 0.2857 & 0.2857 & 0.2857 \end{bmatrix}$$

$$B^k \to \begin{bmatrix} 0.2022 & 0.2022 & 0.2022 \\ 0.3708 & 0.3708 & 0.3708 \\ 0.4270 & 0.4270 & 0.4270 \end{bmatrix}$$

或用有理数形式：

$$A^k \to \begin{bmatrix} 2/7 & 2/7 & 2/7 \\ 3/7 & 3/7 & 3/7 \\ 2/7 & 2/7 & 2/7 \end{bmatrix} \text{且}$$

$$B^k \to \begin{bmatrix} 18/89 & 18/89 & 18/89 \\ 33/89 & 33/89 & 33/89 \\ 38/89 & 38/89 & 38/89 \end{bmatrix}$$

第 3 章

3.1

1. 1　　3. −5　　5. −23　　7. 4

9. 10，从第 3 行开始．

11. −12，从第 1 列或第 4 行开始．

13. 6，从第 2 行或第 2 列开始．

15. 1　　17. −5

19. $ad-bc,\ cb-da$，交换两行改变行列式的符号．

21. $-2,\ (18+12k)-(20+12k)=-2$．一个行的替换不改变一个行列式的值．

23. $-5,\ k(4)-k(2)+k(-7)=-5k$，用一个常数 k 乘一行后的行列式等于 k 乘以原行列式．

25. 1　　27. k　　29. −1

31. 1，矩阵是上三角形或下三角形形式，在对角线上只有 1．行列式等于 1，是对角线元素的乘积．

33. $$\det EA = \det\begin{bmatrix} c & d \\ a & b \end{bmatrix} = cb-ad = (-1)\ (ad-bc)$$
$$= (\det E)(\det A)$$

35. $$\det EA = \det\begin{bmatrix} a+kc & b+kd \\ c & d \end{bmatrix}$$
$$= (a+kc)d-(b+kd)c$$
$$= ad+kcd-bc-kdc = (+1)\ (ad-bc)$$
$$= (\det E)\ (\det A)$$

37. $5A = \begin{bmatrix} 15 & 5 \\ 20 & 10 \end{bmatrix}$；否

39. 在“学习指导”中有提示．

41. 平行四边形的面积和 $[\mathbf{u}\ \ \mathbf{v}]$ 的行列式都等于 6，如果对任何 x，$\mathbf{v}=\begin{bmatrix} x \\ 2 \end{bmatrix}$，面积仍为 6，在每种情形下，平行四边形的底边没有改变，且高度保持为 2，因为 $\mathbf{v}$ 的第二个坐标总是 2．

43. [M]一般地，$\det(A+B)$ 不等于 $\det A+\det B$．

45. [M]在学到 3.2 节时，你可以检查你的猜测．

3.2

1. 交换两行改变行列式的符号．

3. 行替换变换不改变行列式．

5. 3　　7. 0　　9. 3　　11. 120

13. 6　　15. 35　　17. −7　　19. 14

21. 可逆　　23. 不可逆

25. 线性无关　　27. 见“学习指导”

29. −32

31. 提示：证明 $(\det A)\ (\det A^{-1})=1$．

33. 提示：用定理 6．

35. 提示：用定理 6 和另一个定理．

37. $$\det AB = \det\begin{bmatrix} 6 & 0 \\ 17 & 4 \end{bmatrix}$$
$$= 24$$
$$(\det A)\ (\det B) = 3\cdot 8 = 24$$

39. a. −12 b. 500 c. −3 d. $\frac{1}{4}$ e. 64

41. $\det A=(a+e)d-(b+f)c=ad+ed-bc-fc$
$=(ad-bc)+(ed-fc)=\det B+\det C$

43. 提示：用第 3 列的余因子展开计算 $\det A$.

45. [M] 在你做出关于 $A^{\mathrm{T}}A$ 和 AA^{T} 的猜测后参考“学习指导”.

3.3

1. $\begin{bmatrix}5/6\\-1/6\end{bmatrix}$ 3. $\begin{bmatrix}4\\5/2\end{bmatrix}$ 5. $\begin{bmatrix}3/2\\4\\-7/2\end{bmatrix}$

7. $s\neq\pm\sqrt{3}$；$x_1=\dfrac{5s+4}{6(s^2-3)}$，$x_2=\dfrac{-4s-15}{4(s^2-3)}$

9. $s\neq 0,-1$；$x_1=\dfrac{1}{3(s+1)}$，$x_2=\dfrac{4s+3}{6s(s+1)}$

11. $\operatorname{adj}A=\begin{bmatrix}0&1&0\\-3&-1&-3\\3&2&6\end{bmatrix}$，$A^{-1}=\frac{1}{3}\begin{bmatrix}0&1&0\\-3&-1&-3\\3&2&6\end{bmatrix}$

13. $\operatorname{adj}A=\begin{bmatrix}-1&-1&5\\1&-5&1\\1&7&-5\end{bmatrix}$，$A^{-1}=\frac{1}{6}\begin{bmatrix}-1&-1&5\\1&-5&1\\1&7&-5\end{bmatrix}$

15. $\operatorname{adj}A=\begin{bmatrix}2&0&0\\2&6&0\\-1&-9&3\end{bmatrix}$，$A^{-1}=\frac{1}{6}\begin{bmatrix}2&0&0\\2&6&0\\-1&-9&3\end{bmatrix}$

17. 如果 $A=\begin{bmatrix}a&b\\c&d\end{bmatrix}$，那么 $C_{11}=d,C_{12}=-c,C_{21}=-b,C_{22}=a$. 共轭矩阵是余因子的转置：$\operatorname{adj}A=\begin{bmatrix}d&-b\\-c&a\end{bmatrix}$.

根据定理 8，我们用 $\det A$ 来除，得到与 2.2 节中一样的公式.

19. 8 21. 14 23. 22

25. 一个 3×3 矩阵 A 不是可逆的充分必要条件是它的列是线性相关（由矩阵可逆定理）. 这种情形出现的充分必要条件是其中一个列是在其他两列所生成的平面内，这等价于由这些列所确定的平行六面体具有零体积，也就是等价于条件 $\det A=0$.

27. 24 29. $\frac{1}{2}\left|\det[\boldsymbol{v}_1\ \ \boldsymbol{v}_2]\right|$

31. a. 见例 5 b. $4\pi abc/3$

33. [M]在 MATLAB 中，$B-\mathrm{inv}(A)$ 中的元素近似为 10^{-15} 或更小，见“学习指导”的建议，它也许会节省你工作中的键盘输入.

35. [M]MATLAB 学生版 4.0 用 57 771 次浮算来计算 inv(A) 和 14,269 045 次浮算来计算求逆公式. inv(A) 命令仅需要求逆公式中 0.4%的运算量. “学习指导”给出如何使用 flops 命令.

第 3 章补充习题

1. a. T b. T c. F d. F
e. F f. F g. T h. T
i. F j. F k. T l. F
m. F n. T o. F p. T

习题 3 的解基于这样的事实，如果一个矩阵包含两行（或两列）相互为倍数，那么由定理 4，因为矩阵不可逆，矩阵的行列式为零.

3. 做两次行替换变换且提出第 2 行的公共因子，和第 3 行的公共因子.

$$\begin{vmatrix}1&a&b+c\\1&b&a+c\\1&c&a+b\end{vmatrix}=\begin{vmatrix}1&a&b+c\\0&b-a&a-b\\0&c-a&a-c\end{vmatrix}$$
$$=(b-a)\ (c-a)\begin{vmatrix}1&a&b+c\\0&1&-1\\0&1&-1\end{vmatrix}$$
$$=0$$

5. −12

7. 当行列式用第一行的余因子展开，方程具有形式 $ax+by+c=0$，此处 b 和 c 至少有一个非零. 这是一条直线的方程，很明显 (x_1,y_1) 和 (x_2,y_2) 在直线上，因为当用其中一个点的坐标代入 x 和 y 时，行列式的两行相等，于是行列式为零.

9. $T\sim\begin{bmatrix}1&a&a^2\\0&b-a&b^2-a^2\\0&c-a&c^2-a^2\end{bmatrix}$. 因此，由定理 3

$$\det T=(b-a)\ (c-a)\det\begin{bmatrix}1&a&a^2\\0&1&b+a\\0&1&c+a\end{bmatrix}$$
$$=(b-a)\ (c-a)\det\begin{bmatrix}1&a&a^2\\0&1&b+a\\0&0&c-b\end{bmatrix}$$
$$=(b-a)\ (c-a)\ (c-b)$$

11. 面积＝12. 如果从四个顶点减去一个顶点，记新顶点是 $\mathbf{0}, \mathbf{v}_1, \mathbf{v}_2$ 和 $\mathbf{v}_3$，那么变换后的图形（因此原来的图形）是平行四边形的充分必要条件是在 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ 中，某一个是另外两个向量之和.

13. 由求逆公式，$(\text{adj}\ A)\cdot\dfrac{1}{\det A}A=A^{-1}A=I$，由可逆矩阵定理，$\text{adj}A$ 是可逆的，且 $(\text{adj}\ A)^{-1}=\dfrac{1}{\det A}A$.

15. a. $X=CA^{-1}, Y=D-CA^{-1}B$. 现在利用练习 14(c).

b. 从（a）部分和行列式的乘积性质，

$$\det\begin{bmatrix}A&B\\C&D\end{bmatrix}=\det[A(D-CA^{-1}B)]$$
$$=\det[AD-ACA^{-1}B]$$
$$=\det(AD-CAA^{-1}B)$$
$$=\det(AD-CB)$$

其中等式 $AC=CA$ 在第三步中用到.

17. 首先考虑 $n=2$ 的情形，通过直接计算 B 和 C 的行列式来证明公式成立. 现假设该公式对所有 $(k-1)\times(k-1)$ 矩阵成立，设 A,B 和 C 是 $k\times k$ 矩阵. 使用按第一列的余因子展开和归纳假设来求 $\det B$. 对 C 做行变换得到第一个主元下面为零元素的三角形矩阵. 求该矩阵的行列式，然后加到 $\det B$ 中得到最后结果.

19. [M]计算：

$$\begin{vmatrix}1&1&1\\1&2&2\\1&2&3\end{vmatrix}=1,\begin{vmatrix}1&1&1&1\\1&2&2&2\\1&2&3&3\\1&2&3&4\end{vmatrix}=1,\begin{vmatrix}1&1&1&1&1\\1&2&2&2&2\\1&2&3&3&3\\1&2&3&4&4\\1&2&3&4&5\end{vmatrix}=1$$

猜想：
$$\begin{vmatrix}1&1&1&\cdots&1\\1&2&2&&2\\1&2&3&&3\\\vdots&&&\ddots&\vdots\\1&2&3&\cdots&n\end{vmatrix}=1$$

为肯定这个猜想，使用行替换变换产生第一个主元下面的零元素，然后再用第一个主元，如此类推，最终矩阵是 $\begin{vmatrix}1&1&1&\cdots&1\\0&1&1&&1\\0&0&1&&1\\\vdots&&&\ddots&\vdots\\0&0&0&\cdots&1\end{vmatrix}$，是行列式为 1 的上三角形矩阵.

第 4 章

4.1

1. a. $\mathbf{u}+\mathbf{v}$ 在 V 中，因为它的元素都是非负的.

b. 例子:如果 $\mathbf{u}=\begin{bmatrix}2\\2\end{bmatrix}$ 和 $c=-1$，那么 $\mathbf{u}$ 在 V 中，但 $c\mathbf{v}$ 不在 V 中.

3. 例子：如果 $\mathbf{u}=\begin{bmatrix}0.5\\0.5\end{bmatrix}$ 和 $c=4$，那么 $\mathbf{u}$ 在 H 中，但 $c\mathbf{u}$ 不在 H 中.

5. 是，由定理 1，因为集合是 $\text{Span}\{t^2\}$，它是 P_n 的一个子空间.

7. 否，该集合对不是整数的标量乘法不封闭.

9. $H=\text{Span}\{\mathbf{v}\}$，此处 $\mathbf{v}=\begin{bmatrix}-2\\5\\3\end{bmatrix}$.由定理 1，$H$ 是 $\mathbb{R}^3$ 的一个子空间.

11. $W=\text{Span}\{\mathbf{u},\mathbf{v}\}$，此处 $\mathbf{u}=\begin{bmatrix}2\\-1\\0\end{bmatrix}$，$\mathbf{v}=\begin{bmatrix}3\\0\\2\end{bmatrix}$，由定理 1，$W$ 是 $\mathbb{R}^3$ 的一个子空间.

13. a. $\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$ 中仅包含三个向量，且 $\mathbf{w}$ 不在其中.

b. 在 $\text{Span}\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$ 中有无穷多向量.

c. $\mathbf{w}$ 在 $\text{Span}\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$ 中，因为 $\mathbf{w}=\mathbf{v}_1+\mathbf{v}_2$.

15. W 不是向量空间，因为零向量不在 W 中.

17. $S=\left\{\begin{bmatrix}2\\0\\-1\\0\end{bmatrix},\begin{bmatrix}-1\\3\\0\\3\end{bmatrix},\begin{bmatrix}0\\-1\\3\\0\end{bmatrix}\right\}$

19. 提示：利用定理 1.

注意 尽管“学习指导”对每一个单数的这里的答案仅是“提示”的练习有一个完整解答，但你必须自己真正去求这些问题的解 否则，你不能从习题中受益.

21. 是，对一个子空间的条件是明显满足的：零矩阵在 H 中，两个上三角形矩阵的和是上三角形矩阵，且任何一个上三角矩阵的标量乘法仍然是上三角形矩阵.

23. 写出答案后，参考“学习指导” .

25. 4　　27. a. 8　b. 3　c. 5　d. 4

29. $\boldsymbol{u}+(-1)\boldsymbol{u}=1\boldsymbol{u}+(-1)\boldsymbol{u}$　　公理 10

$=[1+(-1)]\boldsymbol{u}$　　公理 8

$=0\boldsymbol{u}=\boldsymbol{0}$　　习题 27

从习题 26，得出 $(-1)\boldsymbol{u}=-\boldsymbol{u}$.

31. 任何包含 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的子空间 H 必须包含 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的所有标量乘法，因而包含 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 所有标量乘法之和，于是 H 必须包含 $\mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$.

33. 提示：对部分解，考虑 $H+K$ 中的 $\boldsymbol{w}_1$ 和 $\boldsymbol{w}_2$，并把 $\boldsymbol{w}_1$ 和 $\boldsymbol{w}_2$ 分别写成 $\boldsymbol{w}_1=\boldsymbol{u}_1+\boldsymbol{v}_1$ 和 $\boldsymbol{w}_2=\boldsymbol{u}_2+\boldsymbol{v}_2$，其中 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 属于 H，而 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 属于 K.

35. [M] 简化的 $[\boldsymbol{v}_1\ \boldsymbol{v}_2\ \boldsymbol{v}_3\ \boldsymbol{w}]$ 阶梯形式表明 $\boldsymbol{w}=\boldsymbol{v}_1-2\boldsymbol{v}_2+\boldsymbol{v}_3$. 因此 $\boldsymbol{w}$ 在由 $\boldsymbol{v}_1$，$\boldsymbol{v}_2$，$\boldsymbol{v}_3$ 生成的子空间中.

37. [M]函数是 $\cos 4t$ 和 $\cos 6t$，见 4.5 节习题 34.

4.2

1. $\begin{bmatrix}3&-5&-3\\6&-2&0\\-8&4&1\end{bmatrix}\begin{bmatrix}1\\3\\-4\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$，所以 $\boldsymbol{w}$ 在 $\mathrm{Nul}\,A$ 中.

3. $\begin{bmatrix}2\\-3\\1\\0\end{bmatrix},\begin{bmatrix}-4\\2\\0\\1\end{bmatrix}$　　5. $\begin{bmatrix}4\\1\\0\\0\\0\end{bmatrix},\begin{bmatrix}-2\\0\\5\\1\\0\end{bmatrix}$

7. W 不是 $\mathbb{R}^3$ 的一个子空间，因为零向量(0,0,0)不属于 W .

9. W 是 $\mathbb{R}^4$ 的一个子空间，由于 W 是下列方程组的解集：

$$\begin{aligned}p-3q-4s\quad\ &=0\\2p\qquad\ -s-5r&=0\end{aligned}$$

11. W 不是子空间，因为 $\boldsymbol{0}$ 不在 W 中，验证：如果一个元素 $(s-2t,3+3s,3s+t,2s)$ 是零，那么 $3+3s=0$ 和 $2s=0$，这是不可能的.

13. $W=\mathrm{Col}\,A$，其中 $A=\begin{bmatrix}1&-6\\0&1\\1&0\end{bmatrix}$，所以由定理 3，$W$ 是一个向量空间.

15. $\begin{bmatrix}0&2&1\\1&-1&2\\3&1&0\\2&-1&-1\end{bmatrix}$

17. a.2　b.4　　19.a.5　b.2

21. 向量 $\begin{bmatrix}2\\3\end{bmatrix}$ 在 $\mathrm{Nul}\,A$ 中，向量 $\begin{bmatrix}6\\-3\\-9\\9\end{bmatrix}$ 在 $\mathrm{Col}\,A$ 中，其余答案可能.

23. $\boldsymbol{w}$ 属于 $\mathrm{Nul}\,A$ 和 $\mathrm{Col}\,A$.

25. 见“学习指导”，现在你应知道如何合理使用它.

27. 令 $\boldsymbol{x}=\begin{bmatrix}3\\2\\-1\end{bmatrix}$ 和 $A=\begin{bmatrix}1&-3&-3\\-2&4&2\\-1&5&7\end{bmatrix}$，那么 $\boldsymbol{x}$ 属于 $\mathrm{Nul}\,A$，由于 $\mathrm{Nul}\,A$ 是 $\mathbb{R}^3$ 的一个子空间，$10\boldsymbol{x}$ 属于 $\mathrm{Nul}\,A$.

29. a. $A\boldsymbol{0}=\boldsymbol{0}$，所以零向量在 $\mathrm{Col}\,A$ 中.

b. 由矩阵乘法的性质，$A\boldsymbol{x}+A\boldsymbol{w}=A(\boldsymbol{x}+\boldsymbol{w})$，它表明 $A\boldsymbol{x}+A\boldsymbol{w}$ 是 A 的列的线性组合，因此属于 $\mathrm{Col}\,A$.

c. $c(A\boldsymbol{x})=A(c\boldsymbol{x})$，这表明对所有数 c，$c(A\boldsymbol{x})$ 在 $\mathrm{Col}\,A$ 中.

31. a. 对任何在 $\mathbb{P}_2$ 中的多项式 $\boldsymbol{p},\boldsymbol{q}$ 和任何数 c，

$$T(\boldsymbol{p}+\boldsymbol{q})=\begin{bmatrix}(\boldsymbol{p}+\boldsymbol{q})\ (0)\\(\boldsymbol{p}+\boldsymbol{q})\ (1)\end{bmatrix}=\begin{bmatrix}\boldsymbol{p}(0)+\boldsymbol{q}(0)\\\boldsymbol{p}(1)+\boldsymbol{q}(1)\end{bmatrix}$$
$$=\begin{bmatrix}\boldsymbol{p}(0)\\\boldsymbol{p}(1)\end{bmatrix}+\begin{bmatrix}\boldsymbol{q}(0)\\\boldsymbol{q}(1)\end{bmatrix}=T(\boldsymbol{p})+T(\boldsymbol{q})$$
$$T(c\boldsymbol{p})=\begin{bmatrix}c\boldsymbol{p}(0)\\c\boldsymbol{p}(1)\end{bmatrix}=c\begin{bmatrix}\boldsymbol{p}(0)\\\boldsymbol{p}(1)\end{bmatrix}=cT(\boldsymbol{p})$$

所以 T 是一个从 $\mathbb{P}_2$ 到 $\mathbb{P}_2$ 的线性变换.

b. 任何在 0 和 1 点为零的二次多项式必须是 $\boldsymbol{p}(t)=t(t-1)$ 的倍数， T 的值域为 $\mathbb{R}^2$.

33\. a. 对属于 $M_{2\times2}$ 中的 A,B 和任何标量 c，

$$\begin{aligned}T(A+B)&=(A+B)+(A+B)^{\mathrm{T}}\\&=A+B+A^{\mathrm{T}}+B^{\mathrm{T}}\qquad\text{转置性质}\\&=(A+A^{\mathrm{T}})+(B+B^{\mathrm{T}})=T(A)+T(B)\end{aligned}$$
$$\begin{aligned}T(cA)&=(cA)+(cA)^{\mathrm{T}}=cA+cA^{\mathrm{T}}\\&=c(A+A^{\mathrm{T}})=cT(A)\end{aligned}$$

所以 T 是一个从 $M_{2\times2}$ 到 $M_{2\times2}$ 的线性变换.

b. 如果 B 是 $M_{2\times2}$ 中任一元素且具有性质 $B^{\mathrm{T}}=B$，若 $A=\frac{1}{2}B$，那么

$$T(A)=\frac{1}{2}B+\left(\frac{1}{2}B\right)^{\mathrm{T}}=\frac{1}{2}B+\frac{1}{2}B=B$$

c. (b)部分说明 T 的值域包含所有满足 $B^{\mathrm{T}}=B$ 的 B，所以，它充分说明任何在 T 值域中的 B 有这个性质，如果 $B=T(A)$，那么由转置性质，$B^{\mathrm{T}}=(A+A^{\mathrm{T}})^{\mathrm{T}}=A^{\mathrm{T}}+A^{\mathrm{TT}}=A^{\mathrm{T}}+A=B$

d. T 的核是 $\left\{\begin{bmatrix}0&b\\-b&0\end{bmatrix}:b\text{是实数}\right\}$

35\. 提示：验证子空间的三个条件，$T(U)$ 中的典型元素具有形式 $T(\boldsymbol{u}_1)$ 和 $T(\boldsymbol{u}_2)$，此处 $\boldsymbol{u}_1,\boldsymbol{u}_2$ 属于 U.

37\. [M] $\boldsymbol{w}$ 属于 Col A，但不属于 Nul A.（解释为什么？）

39\. [M] A 的简化阶梯形式是

$$\begin{bmatrix}1&0&1/3&0&10/3\\0&1&1/3&0&-26/3\\0&0&0&1&-4\\0&0&0&0&0\end{bmatrix}$$

4.3

1\. 是，3×3 矩阵 $A=\begin{bmatrix}1&1&1\\0&1&1\\0&0&1\end{bmatrix}$ 具有 3 主元位置，由可逆矩阵定理，A 是可逆且它的列构成 $\mathbb{R}^3$ 的一个基.（见例 3.）

3\. 这个集合不能构成 $\mathbb{R}^3$ 的基，集合线性相关且不能生成 $\mathbb{R}^3$.

5\. 否，集合是线性相关，因为零向量在集合中，然而，

$$\begin{bmatrix}3&-3&0&0\\-3&7&0&-3\\0&0&0&5\end{bmatrix}\sim\begin{bmatrix}1&-1&0&0\\0&4&0&-3\\0&0&0&5\end{bmatrix}$$

矩阵的每一行有主元，因而它的列生成 $\mathbb{R}^3$.

7\. 否，集合是线性无关，因为它们不是倍数关系（更精确地讲，没有一个向量是其余向量的倍数）. 然而，向量不能生成 $\mathbb{R}^3$，矩阵 $\begin{bmatrix}-2&6\\3&-1\\0&5\end{bmatrix}$ 最多有 2 个主元，因为它仅有 2 列. 所以，并不是每一行有一个主元.

9\. $\begin{bmatrix}2\\-1\\1\\0\end{bmatrix}$

11\. $\begin{bmatrix}3\\1\\0\end{bmatrix},\begin{bmatrix}-2\\0\\1\end{bmatrix}$

13\. Nul A 的基：$\begin{bmatrix}-6\\-5/2\\1\\0\end{bmatrix},\begin{bmatrix}-5\\-3/2\\0\\1\end{bmatrix}$

Col A 的基：$\begin{bmatrix}-2\\2\\-3\end{bmatrix},\begin{bmatrix}4\\-6\\8\end{bmatrix}$

15\. $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_4,\boldsymbol{v}_5\}$

17\. [M] $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_5\}$

19\. 3 个最简单的答案是 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 或 $\{\boldsymbol{v}_1,\boldsymbol{v}_3\}$ 或 $\{\boldsymbol{v}_2,\boldsymbol{v}_3\}$. 其他答案也可能.

21\. 见“学习指导”的提示.

23\. 提示：利用可逆矩阵定理.

25. 否.(为什么集合不是 H 的基？)

27. $\{\cos\omega t, \sin\omega t\}$.

29. 设 A 是 $n\times k$ 矩阵 $[\mathbf{v}_1 \ \cdots \ \mathbf{v}_k]$，由于 A 的列少于行，不可能 A 的每一行有一个主元位置，由 1.4 节的定理 4，A 的列不能生成 $\mathbb{R}^n$，因此不是 $\mathbb{R}^n$ 的基.

31. 提示:如果 $\{\mathbf{v}_1,\cdots,\mathbf{v}_p\}$ 是线性相关，那么存在不全为零的 $c_1,\cdots,c_p$，使得 $c_1\mathbf{v}_1+\cdots+c_p\mathbf{v}_p=\mathbf{0}$．利用这个方程.

33. 任一多项式不是其他多项式的倍数，因此 $\{\mathbf{p}_1,\mathbf{p}_2\}$ 是 $\mathbb{P}_3$ 中的线性无关集.

35. 设 $\{\mathbf{v}_1,\mathbf{v}_3\}$ 是向量空间 V 中的任一线性无关集，并设 $\mathbf{v}_2,\mathbf{v}_4$ 是 $\mathbf{v}_1$ 和 $\mathbf{v}_3$ 的线性组合，则 $\{\mathbf{v}_1,\mathbf{v}_3\}$ 是 $\mathrm{Span}\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4\}$ 的一组基.

37. [M]你可能很聪明且找出特殊 t 值使得(5)中的方程产生几个零，从而得到可以很容易用手工求解的线性方程组，或者，你可利用 t 值，如 $t=0, 0.1, 0.2, \cdots$，产生一个能用矩阵程序来解的线性方程组.

4.4

1. $\begin{bmatrix}3\\-7\end{bmatrix}$ 3. $\begin{bmatrix}-7\\4\\3\end{bmatrix}$ 5. $\begin{bmatrix}2\\-1\end{bmatrix}$ 7. $\begin{bmatrix}-1\\-1\\3\end{bmatrix}$

9. $\begin{bmatrix}1&2\\-3&-5\end{bmatrix}$ 11. $\begin{bmatrix}5\\1\end{bmatrix}$ 13. $\begin{bmatrix}2\\6\\-1\end{bmatrix}$

15. “学习指导”有提示.

17. $\begin{bmatrix}1\\1\end{bmatrix}=5\mathbf{v}_1-2\mathbf{v}_2=10\mathbf{v}_1-3\mathbf{v}_2+\mathbf{v}_3=-\mathbf{v}_2-\mathbf{v}_3$（无穷多解答）.

19. 提示：由假设，零向量可唯一表示成 S 中元素的线性组合.

21. $\begin{bmatrix}9&2\\4&1\end{bmatrix}$

23. 提示：若 $[\mathbf{u}]_B=[\mathbf{w}]_B$ 对 V 中一些 $\mathbf{u}$ 和 $\mathbf{w}$ 成立，且记 $[\mathbf{u}]_B$ 中的元素为 $c_1,\cdots,c_n$.利用 $[\mathbf{u}]_B$ 的定义.

25. 一个可能的方法：首先，说明如果 $\mathbf{u}_1,\cdots,\mathbf{u}_p$ 是线性相关的，那么 $[\mathbf{u}_1]_B,\cdots,[\mathbf{u}_p]_B$ 是线性相关的 .其次，说明如果 $[\mathbf{u}_1]_B,\cdots,[\mathbf{u}_p]_B$ 是线性相关的，那么 $\mathbf{u}_1,\cdots,\mathbf{u}_p$ 是线性相关的，利用习题中显示的两个方程，一个稍微不同的证明在“学习指导”中给出.

27. 线性无关（验证习题 27~34 中的答案）.

29. 线性相关

31. a. 坐标向量 $\begin{bmatrix}1\\-3\\5\end{bmatrix},\begin{bmatrix}-3\\5\\-7\end{bmatrix},\begin{bmatrix}-4\\5\\-6\end{bmatrix},\begin{bmatrix}1\\0\\-1\end{bmatrix}$ 不能生成 $\mathbb{R}^3$. 因为 $\mathbb{R}^3$ 和 $\mathbb{P}_2$ 同构，对应的多项式不能生成 $\mathbb{P}_2$.

b. 坐标向量 $\begin{bmatrix}0\\5\\1\end{bmatrix},\begin{bmatrix}1\\-8\\-2\end{bmatrix},\begin{bmatrix}-3\\4\\2\end{bmatrix},\begin{bmatrix}2\\-3\\0\end{bmatrix}$ 生成 $\mathbb{R}^3$.因为 $\mathbb{R}^3$ 和 $\mathbb{P}_2$ 同构，对应的多项式生成 $\mathbb{P}_2$.

33. [M]坐标向量 $\begin{bmatrix}3\\7\\0\\0\end{bmatrix},\begin{bmatrix}5\\1\\0\\-2\end{bmatrix},\begin{bmatrix}0\\1\\-2\\0\end{bmatrix},\begin{bmatrix}1\\16\\-6\\2\end{bmatrix}$ 是 $\mathbb{R}^4$ 中的线性相关子集.因为 $\mathbb{R}^4$ 和 $\mathbb{P}_3$ 同构，对应的多项式构成 $\mathbb{P}_3$ 的线性相关子集，因此不能构成 $\mathbb{P}_3$ 的基.

35. [M] $[\mathbf{x}]_B=\begin{bmatrix}-5/3\\8/3\end{bmatrix}$ 37. [M] $\begin{bmatrix}1.3\\0\\0.8\end{bmatrix}$

4.5

1. $\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}-2\\1\\3\end{bmatrix}$；维数是 2.

3. $\begin{bmatrix}0\\1\\0\\1\end{bmatrix},\begin{bmatrix}0\\-1\\1\\2\end{bmatrix},\begin{bmatrix}2\\0\\-3\\0\end{bmatrix}$；维数是 3.

5. $\begin{bmatrix}1\\2\\0\\-3\end{bmatrix},\begin{bmatrix}-2\\0\\-2\\0\end{bmatrix},\begin{bmatrix}0\\5\\2\\6\end{bmatrix}$；维数是 3.

7. 没有基；维数为零. 9. 2 11. 3

13. 2，3　15. 2，3　17. 0，3　19. 见“学习指导”.

21. 提示：你只需验证前 4 个埃尔米特多项式是线性无关，为什么？

23. $[\boldsymbol{p}]_B=(3,6,2,1)$

25. 提示：若 S 确实生成 V，且用生成集定理，这会导致一个矛盾，从而证明生成假设是错的.

27. 提示：利用每一个 $\mathbb{P}_n$ 是 $\mathbb{P}$ 的一个子空间的事实.

29. 验证每一个答案：a. 真，b. 真，c. 真.

31. 提示：由于 H 是有限维空间的一个非零子空间，H 是有限维的且有基，如 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$，首先证明 $\{T(\boldsymbol{v}_1),\cdots,T(\boldsymbol{v}_p)\}$ 生成 $T(H)$.

33. [M]a. 一个基是 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{e}_2,\boldsymbol{e}_3\}$，事实上，$\boldsymbol{e}_2,\cdots,\boldsymbol{e}_5$ 中的任何两个向量可将 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 扩展成为 $\mathbb{R}^5$ 的一个基.

4.6

1. rank $A=2$；$\dim\text{Nul}\ A=2$；

Col A 的基为：$\begin{bmatrix}1\\-1\\5\end{bmatrix},\begin{bmatrix}-4\\2\\-6\end{bmatrix}$

Row A 的基为：$(1,0,-1,5),(0,-2,5,-6)$

Nul A 的基为：$\begin{bmatrix}1\\5/2\\1\\0\end{bmatrix},\begin{bmatrix}-5\\-3\\0\\1\end{bmatrix}$

3. rank $A=3$；$\dim\text{Nul}\ A=3$；

Col A 的基：$\begin{bmatrix}2\\-2\\4\\-2\end{bmatrix},\begin{bmatrix}6\\-3\\9\\3\end{bmatrix},\begin{bmatrix}3\\0\\3\\3\end{bmatrix}$

Row A：$(2,6,-6,6,3,6),(0,3,0,3,3,0),(0,0,0,0,3,0)$

Nul A 的基：$\begin{bmatrix}3\\0\\1\\0\\0\\0\end{bmatrix},\begin{bmatrix}0\\-1\\0\\1\\0\\0\end{bmatrix},\begin{bmatrix}-3\\0\\0\\0\\0\\1\end{bmatrix}$

5. 4,3,3

7. 是；不是.由于 Col A 是 $\mathbb{R}^4$ 的一个 4 维子空间，它和 $\mathbb{R}^4$ 一致，它的零空间不是 $\mathbb{R}^3$，由于 Nul A 中的向量有 7 个元素.由秩定理，Nul A 是 $\mathbb{R}^7$ 中的一个三维子空间.

9. 3，否. 注意到 4×6 矩阵的列在 $\mathbb{R}^4$ 中而不是 $\mathbb{R}^3$ 中. Col A 是 $\mathbb{R}^4$ 的三维子空间.

11. 2

13. 5,5. 在这两种情形，主元的个数不能超过列数或行数.

15. 4　　17. 见“学习指导”.

19. 是，在查阅“学习指导”之前，试给一个解释.

21. 否，解释为什么.

23. 是，仅需要 6 个齐次线性方程组成的方程组.

25. 否，解释为什么.

27. Row A 和 Nul A 属于 $\mathbb{R}^n$，Col A 和 Nul A^{T} 属于 $\mathbb{R}^m$，仅有 4 个不同子空间. 因为 Row $A^{\mathrm{T}}=$ Col A 和 Col $A^{\mathrm{T}}=$ Row A.

29. 记住若 Col $A=\mathbb{R}^m$，精确地有 $\dim\text{Col}\ A=m$，或等价地，对所有 $\boldsymbol{b}$ 方程 $A\boldsymbol{x}=\boldsymbol{b}$ 相容. 由习题 28(b)，当 $\dim\text{Nul}\ A^{\mathrm{T}}=0$，精确地有 $\dim\text{Col}\ A=m$，或等价地，方程 $A^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{0}$ 只有平凡解.

31. $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}=\begin{bmatrix}2a&2b&2c\\-3a&-3b&-3c\\5a&5b&5c\end{bmatrix}$，所有列是 $\boldsymbol{u}$ 的倍数，所以，除非 $a=b=c=0$, Col $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ 是一维的.

33. 提示：设 $A=[\boldsymbol{u}\ \ \boldsymbol{u}_2\ \ \boldsymbol{u}_3]$，如果 $\boldsymbol{u}\neq\boldsymbol{0}$，那么 $\boldsymbol{u}$ 是 Col A 的一个基，为什么？

35. [M]提示：见习题 28 和例 4 之前的论述.

4.7

1. a. $\begin{bmatrix}6&9\\-2&-4\end{bmatrix}$　b. $\begin{bmatrix}0\\-2\end{bmatrix}$　　3. (ii)

5. a. $\begin{bmatrix}4&-1&0\\-1&1&1\\0&1&-2\end{bmatrix}$　b. $\begin{bmatrix}8\\2\\2\end{bmatrix}$

7. $\underset{C\leftarrow B}{P}=\begin{bmatrix}-3&1\\-5&2\end{bmatrix}$，$\underset{B\leftarrow C}{P}=\begin{bmatrix}-2&1\\-5&3\end{bmatrix}$

9. $\underset{C\leftarrow B}{P}=\begin{bmatrix}2&3\\0&-1\end{bmatrix}$，$\underset{B\leftarrow C}{P}=\frac{1}{2}\begin{bmatrix}1&3\\0&-2\end{bmatrix}$

11. 见“学习指导”.

13. $\underset{C\leftarrow B}{P}=\begin{bmatrix}1&3&0\\-2&-5&2\\1&4&3\end{bmatrix}$, $[-1+2t]_B=\begin{bmatrix}5\\-2\\1\end{bmatrix}$

15. a. $\mathcal{B}$ 是 V 的一个基.

 b. 坐标变换是一个线性变换.

 c. 一个矩阵和一个向量的乘积.

 d. $\boldsymbol{v}$ 相对于 $\mathcal{B}$ 的坐标向量.

17. a. [M]

$$P^{-1}=\frac{1}{32}\begin{bmatrix}32&0&16&0&12&0&10\\&32&0&24&0&20&0\\&&16&0&16&0&15\\&&&8&0&10&0\\&&&&4&0&6\\&&&&&2&0\\&&&&&&1\end{bmatrix}$$

 b. $\cos^2 t=(1/2)[1+\cos 2t]$

 $\cos^3 t=(1/4)[3\cos t+\cos 3t]$

 $\cos^4 t=(1/8)[3+4\cos 2t+\cos 4t]$

 $\cos^5 t=(1/16)[10\cos t+5\cos 3t+\cos 5t]$

 $\cos^6 t=(1/32)[10+15\cos 2t+6\cos 4t+\cos 6t]$

19. [M]提示：设 $\mathcal{C}$ 是基 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$，那么 P 的列是 $[\boldsymbol{u}_1]_{\mathcal{C}},[\boldsymbol{u}_2]_{\mathcal{C}}$ 和 $[\boldsymbol{u}_3]_{\mathcal{C}}$.利用 $\mathcal{C}-$坐标向量的定义和矩阵代数计算 $\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3$. 解法在“学习指导”中有讨论. 这里有数值解答：

 a. $\boldsymbol{u}_1=\begin{bmatrix}-6\\-5\\21\end{bmatrix}$, $\boldsymbol{u}_2=\begin{bmatrix}-6\\-9\\32\end{bmatrix}$, $\boldsymbol{u}_3=\begin{bmatrix}-5\\0\\3\end{bmatrix}$

 b. $\boldsymbol{w}_1=\begin{bmatrix}28\\-9\\-3\end{bmatrix}$, $\boldsymbol{w}_2=\begin{bmatrix}38\\-13\\2\end{bmatrix}$, $\boldsymbol{w}_3=\begin{bmatrix}21\\-7\\3\end{bmatrix}$

4.8

1. 如果 $y_k=2^k$，那么 $y_{k+1}=2^{k+1}$ 且 $y_{k+2}=2^{k+2}$，将这些公式代入方程左边得到：

$$\begin{aligned}y_{k+2}+2y_{k+1}-8y_k&=2^{k+2}+2\cdot 2^{k+1}-8\cdot 2^k\\&=2^k(2^2+2\cdot 2-8)\\&=2^k(0)=0\ \text{（对所有 }k\text{）}\end{aligned}$$

由于对所有 k，差分方程成立，2^k 是一个解，同样的计算对 $y_k=(-4)^k$ 成立.

3. 信号 2^k 和 $(-4)^k$ 是线性无关，由于没有一个是另外一个的倍数. 例如，不存在数 c 使得 $2^k=c(-4)^k$ 对所有 k 成立. 由定理 17，习题 1 中差分方程的解集 H 是二维的，由 4.5 节的基定理，两个线性无关的信号 2^k 和 $(-4)^k$ 形成 H 的一个基.

5. 如果 $y_k=(-2)^k$，那么对所有 k，有

$$\begin{aligned}y_{k+2}+4y_{k+1}+4y_k&=(-2)^{k+2}+4(-2)^{k+1}+4(-2)^k\\&=(-2)^k[(-2)^2+4(-2)+4]\\&=(-2)^k(0)=0\end{aligned}$$

类似地，如果 $y_k=k(-2)^k$，那么对所有 k，有

$$\begin{aligned}&y_{k+2}+4y_{k+1}+4y_k\\&=(k+2)(-2)^{k+2}+4(k+1)(-2)^{k+1}+4k(-2)^k\\&=(-2)^k[(k+2)(-2)^2+4(k+1)(-2)+4k]\\&=(-2)^k[4k+8-8k-8+4k]\\&=(-2)^k(0)\end{aligned}$$

这样，$(-2)^k$ 和 $k(-2)^k$ 都是在差分方程的解空间 H 中. 而且，没有数 c，使得 $k(-2)^k=c(-2)^k$ 对所有 k 成立，因为 c 的选取必须与 k 无关.同样，没有数 c 使得 $(-2)^k=ck(-2)^k$ 对所有 k 成立，所以两个信号线性无关. 由于 $\dim H=2$，由基定理，信号形成 H 的一个基.

7. 是 9. 是

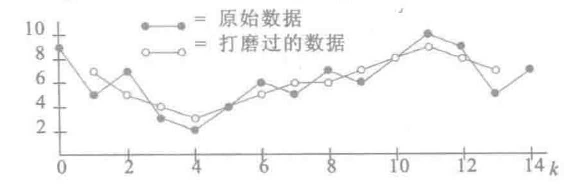
11. 否，两个信号不能生成三维解空间.

13. $\left(\frac{1}{3}\right)^k$, $\left(\frac{2}{3}\right)^k$ 15. $\left(\frac{1}{2}\right)^k$, $\left(\frac{2}{3}\right)^k$

17. $Y_k=c_1(0.8)^k+c_2(0.5)^k+10\to 10$，当 $k\to\infty$ 时

19. $y_k=c_1(-2+\sqrt{3})^k+c_2(-2-\sqrt{3})^k$

21. 7,5,4,3,4,5,6,6,7,8,9,8,7,见下图.



23. a. $y_{k+1}-1.01y_k=-450$，$y_0=10\,000$

25. $k^2+c_1\cdot(-4)^k+c_2$

27. $-2+k+c_1\cdot 2^k+c_2\cdot(-2)^k$

29. $x_{k+1}=Ax_k$，此处

$$A=\begin{bmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\2&-6&8&-3\end{bmatrix},\ x=\begin{bmatrix}y_k\\y_{k+1}\\y_{k+2}\\y_{k+3}\end{bmatrix}$$

31. 方程对所有 k 成立，所以其中 k 用 $(k-1)$ 替换，可将原方程变为：

$y_{k+2}+5y_{k+1}+6y_k=0$ 对所有 k 成立

方程是 2 阶的.

33. 对所有 k，Casorati 矩阵 $C(k)$ 不可逆，在这种情形，Casorati 矩阵没有给出信号集合线性无关和线性相关的信息. 事实上，没有一个信号是其他信号的倍数，所以它们线性无关.

35. 提示：验证定义线性变换的两个性质，对 S 中的 $\{y_k\}$ 和 $\{z_k\}$，研究 $T(\{y_k\}+\{z_k\})$，如果 r 是任意数，那么 $r\{y_k\}$ 的第 k 项是 ry_k；因此 $T(r\{y_k\})$ 是序列 $\{w_k\}$，由下式给出：

$$w_k=ry_{k+2}+a(ry_{k+1})+b(ry_k)$$

37. 提示：求 $TD(y_0,y_1,y_2,\cdots)$和 $DT(y_0,y_1,y_2,\cdots)$.

4.9

1. a. 从 新闻 音乐 到 $\begin{bmatrix}0.7&0.6\\0.3&0.4\end{bmatrix}$ 新闻 音乐　b. $\begin{bmatrix}1\\0\end{bmatrix}$　c. 33%

3. a. 从 健康 疾病 到 $\begin{bmatrix}0.95&0.45\\0.05&0.55\end{bmatrix}$ 健康 疾病　b. 15%, 12.5%

c. 0.925；使用 $x_0=\begin{bmatrix}1\\0\end{bmatrix}$.

5. $\begin{bmatrix}5/14\\9/14\end{bmatrix}$　7. $\begin{bmatrix}1/4\\1/2\\1/4\end{bmatrix}$

9. 是，由于 P^2 都具有正元素.

11. a. $\begin{bmatrix}2/3\\1/3\end{bmatrix}$　b. 2/3

13. a. $\begin{bmatrix}0.9\\0.1\end{bmatrix}$　b. 0.10，否

15. [M]大约 13.9%的美国人口.

17. a. P 的一个列的元素之和为 1，矩阵 $P-I$ 中一个列具有与 P 同样的元素，除了其中一个元素减了 1，因此每一列的和为 0.

b. 由(a)，$P-I$ 的最后一行是其他行之和的相反数.

c. 由(b)和生成定理，$P-I$ 的最后一行可以删去，而剩下的 $(n-1)$ 行，仍然生成行空间. 另一种方法，利用(a)和行变换不改变行空间的事实，取 A 是 $P-I$ 将其他行加到最后一行的矩阵，由(a)，行空间可由 A 的前 $(n-1)$ 行所生成.

d. 由秩定理和(c)，$P-I$ 的列空间的维数小于 n，因而它的零空间是非平凡的.你可以不用秩定理，而用可逆矩阵定理，因为 $P-I$ 是一个方阵.

19. a. 乘积 Sx 等于 x 中元素之和，对一个概率向量，这个和一定等于 1.

b. $P=[p_1\ \ p_2\ \cdots\ p_n]$，此处 p_i 为概率向量，由矩阵乘积和(a),

$SP=[Sp_1\ \ Sp_2\ \cdots\ Sp_n]=[1\ 1\ \cdots\ 1]=S$

c. 由(b)，$S(Px)=(SP)x=Sx=1$，同样 Px 中的元素是非负的(由于 P 和 x 具有非负元素)，因此，由(a)，Px 是一个概率向量.

第 4 章补充习题

1. a. T　b. T　c. F　d. F　e. T
f. T　g. F　h. F　i. T　j. F
k. F　l. F　m. T　n. F　o. T
p. T　q. F　r. T　s. T　t. F

3. 所有 (b_1,b_2,b_3) 的集合，满足 $b_1+2b_2+3b_3=0$.

5. 向量 p_1 不是零，p_2 不是 p_1 的倍数，因此保留这两个向量. 由于 $p_3=2p_1+2p_2$，排除 p_3.由于 p_4 有 t^2 项，它不可能是 p_1,p_2 的线性组合，故保留 p_4. 最后 $p_5=p_1+p_4$，故排除 p_5. 结果基是 $\{p_1,p_2,p_4\}$.

7. 齐次方程组的解集是由两个解生成的，在这种情形下，18×20 的系数矩阵 A 的零空间最多是 2 维. 由秩定理 $\dim\mathrm{Col}\,A\geqslant 20-2=18$，这说明 $\mathrm{Col}\,A=\mathbb{R}^{18}$，因为 A 有 18 行且每一个方程

$Ax=b$ 是相容的.

9. 设 A 是线性变换 T 的标准的 $m\times n$ 矩阵.

   a. 如果 T 是一对一映射，则 A 的列向量线性无关(1.9 节定理 12). 因此 $\dim \operatorname{Nul} A=0$.由秩定理， $\dim \operatorname{Col} A=\operatorname{rank} A=n$.因为 T 的值域是 $\operatorname{Col} A$ ， T 的值域的维数是 n .

   b. 如果 T 是满射，则 A 的列向量生成 $\mathbb{R}^m$ (1.9 节定理 12)，所以 $\dim \operatorname{Col} A=m$.由秩定理， $\dim \operatorname{Nul} A=n-\dim \operatorname{Col} A=n-m$.因为 T 的核是 $\operatorname{Nul} A$ ， T 的核的维数是 $n-m$.

11. 如果 S 是 V 中有限的生成集，那么 S 的一个子集，如 S'，是 V 的一个基.由于 S' 一定生成 V ， S' 不能是 S 的一个真子集，因为 S 是最小的，这样， $S'=S$ ，从而证明 S 是 V 的一个基.

12. a. 提示：对某个 $\boldsymbol{x}$, $\operatorname{Col} AB$ 中的任意 $\boldsymbol{y}$ 有形式 $\boldsymbol{y}=AB\boldsymbol{x}$.

13. 由习题 9， $\operatorname{rank}\ PA\leqslant\operatorname{rank}\ A$ ，且 $\operatorname{rank} A=\operatorname{rank}\ P^{-1}(PA)\leqslant\operatorname{rank}\ PA$ ， 于是 $\operatorname{rank}\ PA=\operatorname{rank}\ A$.

15. 方程 $AB=0$ 说明 B 中的每一列属于 $\operatorname{Nul} A$.由于 $\operatorname{Nul} A$ 是子空间， B 的列的所有线性组合都属于 $\operatorname{Nul} A$ ，因此 $\operatorname{Col} B$ 是 $\operatorname{Nul} A$ 的子空间. 由 4.5 节定理 11， $\dim \operatorname{Col} B\leqslant\dim \operatorname{Nul} A$. 应用秩定理，得到
   $$n=\operatorname{rank} A+\dim \operatorname{Nul} A\geqslant\operatorname{rank} A+\operatorname{rank} B$$

17. a. 设 A_1 由 A 的 r 个主元列组成，A_1 的列是线性无关的，所以， A_1 是秩为 r 的 $m\times r$ 阵.

   b. 将秩定理应用于 A_1， $\operatorname{Row}\ A_1$ 的维数是 r ，所以 A_1 有 r 个线性无关的行. 利用这些行构造 A_2，那么 A_2 是 $r\times r$ 的，且有线性无关的行. 由可逆矩阵定理， A_2 是可逆的.

19. $$[B\ \ AB\ \ A^2B]=\begin{bmatrix}0&1&0\\1&-0.9&0.81\\1&0.5&0.25\end{bmatrix}\sim\begin{bmatrix}1&-0.9&0.81\\0&1&0\\0&0&-0.56\end{bmatrix}$$

   这个矩阵的秩为 3，所以 (A,B) 是可控制的.

21. [M] $\operatorname{rank}(B\ \ AB\ \ A^2B\ \ A^3B)=3$, (A,B) 不可控制.

第 5 章

5.1

1. 是 3. 是，$\lambda=-2$ 5. 是，$\lambda=-5$ 7. 是，$\begin{bmatrix}1\\1\\-1\end{bmatrix}$

9. $\lambda=1:\begin{bmatrix}0\\1\end{bmatrix};\lambda=3:\begin{bmatrix}1\\1\end{bmatrix}$ 11. $\lambda=-1:\begin{bmatrix}3\\2\end{bmatrix};\lambda=7:\begin{bmatrix}-1\\2\end{bmatrix}$

13. $\lambda=1:\begin{bmatrix}0\\1\\0\end{bmatrix}$; $\lambda=2:\begin{bmatrix}-1\\2\\2\end{bmatrix}$; $\lambda=3:\begin{bmatrix}-1\\1\\1\end{bmatrix}$

15. $\begin{bmatrix}-1\\1\\0\end{bmatrix},\begin{bmatrix}-1\\0\\1\end{bmatrix}$ 17. 0，3，−2

19. 0. 验证你的答案.

21. 写出答案之后，见“学习指导”.

23. 提示：用定理 2.

25. 提示：利用方程 $A\boldsymbol{x}=\lambda\boldsymbol{x}$ 找出包含 A^{-1} 的方程.

27. 提示：对任何 λ, $(A-\lambda I)^{\mathrm{T}}=A^{\mathrm{T}}-\lambda I$ ， 由定理（哪一个？） $A^{\mathrm{T}}-\lambda I$ 可逆的充分必要条件是 $A-\lambda I$ 可逆.

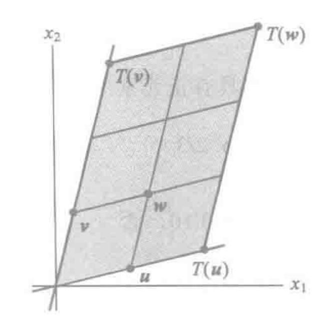
29. 设 $\boldsymbol{v}$ 是 $\mathbb{R}^n$ 中向量，其元素全是 1，那么 $A\boldsymbol{v}=s\boldsymbol{v}$.

31. 提示：如果 A 是 T 的标准矩阵，找一个非零向量 $\boldsymbol{v}$ (平面上一点)使得 $A\boldsymbol{v}=\boldsymbol{v}$.

33. a. $\boldsymbol{x}_{k+1}=c_1\lambda^{k+1}\boldsymbol{u}+c_2\mu^{k+1}\boldsymbol{v}$

   b. $$\begin{aligned}A\boldsymbol{x}_k&=A(c_1\lambda^k\boldsymbol{u}+c_2\mu^k\boldsymbol{v})\\&=c_1\lambda^kA\boldsymbol{u}+c_2\mu^kA\boldsymbol{v}&&\text{线性性质}\\&=c_1\lambda^k\lambda\boldsymbol{u}+c_2\mu^k\mu\boldsymbol{v}&&\boldsymbol{u}\text{ 和 }\boldsymbol{v}\text{ 是特征向量}\\&=\boldsymbol{x}_{k+1}\end{aligned}$$

35. [M]



37. [M] $\lambda=5:\begin{bmatrix}-1\\-1\\2\end{bmatrix}$；$\lambda=10:\begin{bmatrix}-3\\2\\1\end{bmatrix}$；$\lambda=15:\begin{bmatrix}2\\2\\1\end{bmatrix}$.

39. [M] $\lambda=-4:\begin{bmatrix}0\\1\\2\\0\\1\end{bmatrix}$；$\lambda=-8:\begin{bmatrix}6\\3\\3\\2\\0\end{bmatrix},\begin{bmatrix}0\\0\\-1\\0\\1\end{bmatrix}$；

$$\lambda=12:\begin{bmatrix}0\\0\\-1\\1\\0\end{bmatrix},\begin{bmatrix}2\\1\\2\\0\\1\end{bmatrix}$$

5.2

1. $\lambda^2-4\lambda-45$；9, −5　　3. $\lambda^2-3\lambda-40$；−5,8

5. $\lambda^2-16\lambda+48$；4,12

7. $\lambda^2-9\lambda+32$；无实特征值

9. $-\lambda^3+10\lambda^2-33\lambda+36$　　11. $-\lambda^3+8\lambda^2-19\lambda+12$

13. $-\lambda^3+18\lambda^2-95\lambda+150$　　15. 2，3，5，5

17. 3，3，1，1，0

19. 提示：方程对所有 λ 成立.

21. 见“学习指导”的提示.

23. 提示：求一个可逆矩阵 P，使得 $RQ=P^{-1}AP$.

25. a. $\{v_1,v_2\}$，此处 $v_2=\begin{bmatrix}-1\\1\end{bmatrix}$是对应 $\lambda=0.3$ 的特征向量.

b. $x_0=v_1-\frac{1}{14}v_2$.

c. $x_1=v_1-\frac{1}{14}(0.3)v_2$，$x_2=v_1-\frac{1}{14}(0.3)^2v_2$，$x_k=v_1-\frac{1}{14}(0.3)^kv_2$.当 $k\to\infty$时，$(0.3)^k\to0$，所以 $x_k\to v_1$.

27. a. $Av_1=v_1, Av_2=0.5v_2, Av_3=0.2v_3$.（这也证明 A 的特征值为 1，0.5 和 0.2.）

b. $\{v_1,v_2,v_3\}$ 线性无关，因为特征向量对应于不同的特征值（定理 2）.由于集合中包含三个向量，集合是 $\mathbb{R}^3$ 的一个基.所以存在(唯一)常数使得 $x_0=c_1v_1+c_2v_2+c_3v_3$，那么

$$w^Tx_0=c_1w^Tv_1+c_2w^Tv_2+c_3w^Tv_3 \qquad (*)$$

由于 x_0 和 v_1 是概率向量，且 v_2 和 v_3 的元素之和为 0，(*)表明 $1=c_1$.

c. 由(b), $x_0=v_1+c_2v_2+c_3v_3$. 利用(a),

$$\begin{aligned}x_k=A^kx_0&=A^kv_1+c_2A^kv_2+c_3A^kv_3\\&=v_1+c_2(0.5)^kv_2+c_3(0.2)^kv_3\\&\to v_1,\text{当}k\to\infty\end{aligned}$$

29. [M]报告你的结果和结论，如果使用“学习指导”中讨论的 gauss 程序，你可避免复杂计算.

5.3

1. $\begin{bmatrix}226&-525\\90&-209\end{bmatrix}$　　3. $\begin{bmatrix}a^k&0\\2(a^k-b^k)&b^k\end{bmatrix}$

5. $\lambda=2:\begin{bmatrix}-1\\1\\-1\end{bmatrix}$；$\lambda=3:\begin{bmatrix}1\\-1\\0\end{bmatrix},\begin{bmatrix}0\\-1\\1\end{bmatrix}$

当一个答案包含一个对角化时，$A=PDP^{-1}$，因子 P 和 D 不唯一，所以你的答案也许与这里给出的不同.

7. $P=\begin{bmatrix}1&0\\3&1\end{bmatrix}$, $D=\begin{bmatrix}1&0\\0&-1\end{bmatrix}$

9. 不能对角化

11. $P=\begin{bmatrix}1&-1&-1\\2&1&0\\3&0&1\end{bmatrix}$, $D=\begin{bmatrix}5&0&0\\0&-1&0\\0&0&-1\end{bmatrix}$

13. $P=\begin{bmatrix}-1&2&1\\-1&-1&0\\1&0&1\end{bmatrix}$, $D=\begin{bmatrix}5&0&0\\0&1&0\\0&0&1\end{bmatrix}$

15. $P=\begin{bmatrix}-1&-1&-1\\1&1&0\\-1&0&1\end{bmatrix}$, $D=\begin{bmatrix}0&0&0\\0&1&0\\0&0&1\end{bmatrix}$

17. 不能对角化.

19. $P=\begin{bmatrix}1&3&-1&-1\\0&2&-1&2\\0&0&1&0\\0&0&0&1\end{bmatrix}$, $D=\begin{bmatrix}5&0&0&0\\0&3&0&0\\0&0&2&0\\0&0&0&2\end{bmatrix}$

21. 见“学习指导”

23. 是. (解释为什么？)

25. 否，A 一定可以对角化.(解释为什么？)

27. 提示：写 $A=PDP^{-1}$，由于 A 是可逆，0 不是 A 的特征值，所以，D 在对角线上有非零元素.

29. 一个答案是 $P_1=\begin{bmatrix}1 & 1\\ -2 & -1\end{bmatrix}$，它的列是对应 D_1 特征值的特征向量.

31. 提示：构造一个合适的 2×2 三角形矩阵.

33. [M]

$$P=\begin{bmatrix}2 & 0 & -1 & 1\\ 7 & -1 & 0 & -4\\ 7 & 0 & 2 & 0\\ 2 & 1 & 0 & 3\end{bmatrix},\ D=\begin{bmatrix}-12 & 0 & 0 & 0\\ 0 & -12 & 0 & 0\\ 0 & 0 & 13 & 0\\ 0 & 0 & 0 & 13\end{bmatrix}$$

35. [M]

$$P=\begin{bmatrix}2 & 1 & -3 & 1 & 0\\ 1 & 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0 & -1\\ 0 & 1 & 0 & 1 & 0\\ 0 & 0 & 2 & 0 & 1\end{bmatrix},$$

$$D=\begin{bmatrix}7 & 0 & 0 & 0 & 0\\ 0 & 7 & 0 & 0 & 0\\ 0 & 0 & 7 & 0 & 0\\ 0 & 0 & 0 & -14 & 0\\ 0 & 0 & 0 & 0 & -14\end{bmatrix}$$

5.4

1. $\begin{bmatrix}3 & -1 & 0\\ -5 & 6 & 4\end{bmatrix}$

3. a. $T(\boldsymbol{e}_1)=\boldsymbol{b}_3$，$T(\boldsymbol{e}_2)=-\boldsymbol{b}_1-2\boldsymbol{b}_2$，$T(\boldsymbol{e}_3)=2\boldsymbol{b}_1+3\boldsymbol{b}_3$

b. $[T(\boldsymbol{e}_1)]_B=\begin{bmatrix}0\\0\\1\end{bmatrix}$，$[T(\boldsymbol{e}_2)]_B=\begin{bmatrix}-1\\-2\\0\end{bmatrix}$，

$[T(\boldsymbol{e}_3)]_B=\begin{bmatrix}2\\0\\3\end{bmatrix}$

c. $\begin{bmatrix}0 & -1 & 2\\ 0 & -2 & 0\\ 1 & 0 & 3\end{bmatrix}$

5. a. $9-3t+t^2+t^3$

b. 对任意属于 $\mathbb{P}_2$ 的 $\boldsymbol{p},\boldsymbol{q}$ 和任意数 c，

$$\begin{aligned}T[\boldsymbol{p}(t)+\boldsymbol{q}(t)]&=(t+3)[\boldsymbol{p}(t)+\boldsymbol{q}(t)]\\&=(t+3)\boldsymbol{p}(t)+(t+3)\boldsymbol{q}(t)\\&=T[\boldsymbol{p}(t)]+T[\boldsymbol{q}(t)]\\T[c\cdot\boldsymbol{p}(t)]&=(t+3)[c\cdot\boldsymbol{p}(t)]=c\cdot(t+3)\boldsymbol{p}(t)\\&=c\cdot T[\boldsymbol{p}(t)]\end{aligned}$$

c. $\begin{bmatrix}3 & 0 & 0\\ 1 & 3 & 0\\ 0 & 1 & 3\\ 0 & 0 & 1\end{bmatrix}$

7. $\begin{bmatrix}3 & 0 & 0\\ 5 & -2 & 0\\ 0 & 4 & 1\end{bmatrix}$

9. a. $\begin{bmatrix}2\\5\\8\end{bmatrix}$

b. 提示：对任意 $\mathbb{P}_2$ 中的 $\boldsymbol{p},\boldsymbol{q}$ 和任意数 c，计算 $T(\boldsymbol{p}+\boldsymbol{q})$ 和 $T(c\cdot\boldsymbol{p})$.

c. $\begin{bmatrix}1 & -1 & 1\\ 1 & 0 & 0\\ 1 & 1 & 1\end{bmatrix}$

11. $\begin{bmatrix}-2 & -2\\ 0 & -1\end{bmatrix}$

13. $\boldsymbol{b}_1=\begin{bmatrix}1\\1\end{bmatrix}$，$\boldsymbol{b}_2=\begin{bmatrix}1\\3\end{bmatrix}$

15. $\boldsymbol{b}_1=\begin{bmatrix}2\\1\end{bmatrix}$，$\boldsymbol{b}_2=\begin{bmatrix}-1\\3\end{bmatrix}$

17. a. $A\boldsymbol{b}_1=3\boldsymbol{b}_1$，所以 $\boldsymbol{b}_1$ 是 A 的特征向量，然而，A 仅有一个特征值，$\lambda=3$，且特征子空间仅是一维的，所以，A不可对角化.

b. $\begin{bmatrix}3 & 1\\ 0 & 3\end{bmatrix}$

19. 由定义，如果 A 与 B 相似，存在一个可逆矩阵 P，使得 $P^{-1}AP=B$.(见 5.2 节.)那么 B 是可逆的，因为它是可逆矩阵的乘积. 为证明 A^{-1} 与 B^{-1} 相似，利用方程 $P^{-1}AP=B$. 见"学习指导".

21. 提示：复习练习题 2.

23. 提示：计算 $B(P^{-1}\boldsymbol{x})$.

25. 提示：写出 $A=PBP^{-1}=(PB)P^{-1}$，且利用迹的性质.

27. 对每一个 j，$I(\boldsymbol{b}_j)=\boldsymbol{b}_j$，由于 $\mathbb{R}^n$ 中任意向量的

标准坐标向量恰好是向量自身，$[I(\boldsymbol{b}_j)]_\varepsilon=\boldsymbol{b}_j$，$I$ 相对于 $\mathcal{B}$ 和标准基 $\mathcal{E}$ 的矩阵简单表示为 $[\boldsymbol{b}_1\ \boldsymbol{b}_2\ \cdots\ \boldsymbol{b}_n]$.这个矩阵就是 4.4 节定义的坐标变换矩阵 $P_\mathcal{B}$.

29. 恒等变换的 $\mathcal{B}$ 矩阵是 I_n，因为第 j 个基向量 $\boldsymbol{b}_j$ 的 $\mathcal{B}$ 坐标向量是 I_n 的第 j 列.

31. [M] $\begin{bmatrix}-7&-2&-6\\0&-4&-6\\0&0&-1\end{bmatrix}$

5.5

1. $\lambda=2+\mathrm{i}$, $\begin{bmatrix}-1+\mathrm{i}\\1\end{bmatrix}$; $\lambda=2-\mathrm{i}$, $\begin{bmatrix}-1-\mathrm{i}\\1\end{bmatrix}$

3. $\lambda=3+2\mathrm{i}$, $\begin{bmatrix}-1-\mathrm{i}\\4\end{bmatrix}$; $\lambda=3-2\mathrm{i}$, $\begin{bmatrix}-1+\mathrm{i}\\4\end{bmatrix}$

5. $\lambda=4+\mathrm{i}$, $\begin{bmatrix}1-\mathrm{i}\\2\end{bmatrix}$; $\lambda=4-\mathrm{i}$, $\begin{bmatrix}1+\mathrm{i}\\2\end{bmatrix}$

7. $\lambda=\sqrt{3}\pm\mathrm{i},\varphi=\pi/6$ 弧度，$r=2$

9. $\lambda=\pm2\mathrm{i},\varphi=-\pi/2$ 弧度，$r=2$

11. $\lambda=-\sqrt{3}\pm\mathrm{i},\varphi=-5\pi/6$ 弧度，$r=2$

在习题 13~20 中，其他答案也有可能.任何 P，使得 $P^{-1}AP$ 等于给定的 C 或 C^{T} 的都是可满足的答案.首先求 P；然后计算 $P^{-1}AP$.

13. $P=\begin{bmatrix}-1&-1\\1&0\end{bmatrix}$, $C=\begin{bmatrix}2&-1\\1&2\end{bmatrix}$

15. $P=\begin{bmatrix}3&1\\0&2\end{bmatrix}$, $C=\begin{bmatrix}1&3\\-3&1\end{bmatrix}$

17. $P=\begin{bmatrix}1&2\\0&-5\end{bmatrix}$, $C=\begin{bmatrix}-3&4\\-4&-3\end{bmatrix}$

19. $P=\begin{bmatrix}2&-1\\2&0\end{bmatrix}$, $C=\begin{bmatrix}0.96&-0.28\\0.28&0.96\end{bmatrix}$

21. $\boldsymbol{y}=\begin{bmatrix}2\\-1+2\mathrm{i}\end{bmatrix}=\dfrac{-1+2\mathrm{i}}{5}\begin{bmatrix}-2-4\mathrm{i}\\5\end{bmatrix}$

23. (a) 共轭的性质和等式 $\overline{\boldsymbol{x}}^{\mathrm{T}}=\overline{\boldsymbol{x}^{\mathrm{T}}}$.
   (b) $\overline{A\boldsymbol{x}}=A\overline{\boldsymbol{x}}$ 且 A 是实的；
   (c) 由于 $\boldsymbol{x}^{\mathrm{T}}A\overline{\boldsymbol{x}}$ 是一个数，因而可认为是一个 1×1 矩阵.
   (d) 转置的性质.
   (e) $A^{\mathrm{T}}=A$，q 的定义.

25. 提示：首先写出 $\boldsymbol{x}=\mathrm{Re}\,\boldsymbol{x}+i(\mathrm{Im}\,\boldsymbol{x})$

27. [M] $P=\begin{bmatrix}-1&1&-1&-1\\0&-1&0&2\\1&0&0&-2\\0&0&2&0\end{bmatrix}$,

$C=\begin{bmatrix}-2&5&0&0\\-5&-2&0&0\\0&0&-4&10\\0&0&-10&-4\end{bmatrix}$

其他选择也是可能的，但 C 必须等于 $P^{-1}AP$.

5.6

1. a. 提示：求 c_1,c_2，使得 $\boldsymbol{x}_0=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2$，利用这个表达式和事实 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 是 A 的特征向量，计算 $\boldsymbol{x}_1=\begin{bmatrix}49/3\\41/3\end{bmatrix}$.
   b. 一般地，$\boldsymbol{x}_k=5(3)^k\boldsymbol{v}_1-4\left(\dfrac{1}{3}\right)^k\boldsymbol{v}_2$，对 $k\geqslant0$.

3. 当 $p=0.2$，A 的特征值为 0.9 和 0.7，且

$\boldsymbol{x}_k=c_1(0.9)^k\begin{bmatrix}1\\1\end{bmatrix}+c_2(0.7)^k\begin{bmatrix}2\\1\end{bmatrix}\to\boldsymbol{0}$，当 $k\to\infty$ 时

捕食者的出生率越高，猫头鹰的食物供应就越少，最后捕食者和被捕食者的数量都会消失.

5. 如果 $p=0.325$，特征值是 1.05 和 0.55，由于 1.05>1，两种数量每年会增加 5%，对应 1.05 的特征向量为(6,13)，所以，最后每 13000 只松鼠大约有 6 只斑点猫头鹰.

7. a. 原点是鞍点，因为 A 的特征值的绝对值一个大于 1，另一个小于 1.
   b. 最大的吸引方向是特征值 1/3 对应的特征向量 $\boldsymbol{v}_2$，所有是 $\boldsymbol{v}_2$ 倍数的向量都被吸引到原点，最大的排斥方向由特征向量 $\boldsymbol{v}_1$ 给出，所有 $\boldsymbol{v}_1$ 的倍数都被排斥.
   c. 见“学习指导”.

9. 鞍点，A 的特征值 2，0.5；最大的排斥方向：通过(0,0)和(−1,1)的直线.最大吸引方向，通过(0,0)和(1,4)的直线.

11. 吸引子；特征值：0.9，0.8；最大吸引方向：通过(0,0)和(5,4)的直线.

13. 排斥子：特征值 1.2，1.1；最大排斥方向，通过(0,0)和(3,4)的直线.

15. $x_k = v_1 + (0.1)(0.5)^k \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} + (0.3)(0.2)^k \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \to v_1$ 当 $k \to \infty$.

17. a. $A = \begin{bmatrix} 0 & 1.6 \\ 0.3 & 0.8 \end{bmatrix}$

b. 数量会增加，因为 A 的最大特征值是 1.2，此值数值上大于 1.最后的出生增长速度为 1.2，即每年增加 20%.对应特征值 $\lambda_1 = 1.2$ 的特征向量是(4,3)，说明每 3 个成人将有 4 个儿童.

c. [M]5 或 6 年以后，儿童—成人的比率会变得稳定，"学习指导"描述如何构造一个矩阵程序去生成一个数据矩阵，它的列给出每年儿童和成人的数量，数据的图像也有讨论.

5.7

1. $x(t) = \frac{5}{2}\begin{bmatrix} -3 \\ 1 \end{bmatrix} e^{4t} - \frac{3}{2}\begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{2t}$

3. $-\frac{5}{2}\begin{bmatrix} -3 \\ 1 \end{bmatrix} e^{t} + \frac{9}{2}\begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-t}$，原点是鞍点，最大吸引方向是通过(−1,1)和原点的直线.最大排斥方向是通过(−3,1)和原点的直线.

5. $-\frac{1}{2}\begin{bmatrix} 1 \\ 3 \end{bmatrix} e^{4t} + \frac{7}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{6t}$，原点是一个排斥子，最大排斥方向是通过(1,1)和原点的直线.

7. 令 $P = \begin{bmatrix} 1 & 1 \\ 3 & 1 \end{bmatrix}$ 和 $D = \begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix}$，那么 $A = PDP^{-1}$，将 $x = Py$ 代入 $x' = Ax$，我们有

$$\frac{d}{dt}(Py) = A(Py)$$

$$Py' = PDP^{-1}(Py) = PDy$$

用 P^{-1} 左乘，得到 $y' = Dy$ 或

$$\begin{bmatrix} y_1'(t) \\ y_2'(t) \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}$$

9. （复数解）：

$$c_1 \begin{bmatrix} 1-i \\ 1 \end{bmatrix} e^{(-2+i)t} + c_2 \begin{bmatrix} 1+i \\ 1 \end{bmatrix} e^{(-2-i)t}$$

（实数解）：

$$c_1 \begin{bmatrix} \cos t + \sin t \\ \cos t \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} \sin t - \cos t \\ \sin t \end{bmatrix} e^{-2t}$$

轨迹是螺旋趋于原点.

11. （复数解）：$c_1 \begin{bmatrix} -3+3i \\ 2 \end{bmatrix} e^{3it} + c_2 \begin{bmatrix} -3-3i \\ 2 \end{bmatrix} e^{-3it}$

（实数解）：

$$c_1 \begin{bmatrix} -3\cos 3t - 3\sin 3t \\ 2\cos 3t \end{bmatrix} + c_2 \begin{bmatrix} -3\sin 3t + 3\cos 3t \\ 2\sin 3t \end{bmatrix}$$

轨迹是绕原点的椭圆.

13. （复数解）：$c_1 \begin{bmatrix} 1+i \\ 2 \end{bmatrix} e^{(1+3i)t} + c_2 \begin{bmatrix} 1-i \\ 2 \end{bmatrix} e^{(1-3i)t}$

（实数解）：

$$c_1 \begin{bmatrix} \cos 3t - \sin 3t \\ 2\cos 3t \end{bmatrix} e^{t} + c_2 \begin{bmatrix} \sin 3t + \cos 3t \\ 2\sin 3t \end{bmatrix} e^{t}$$

轨迹是螺旋远离原点.

15. [M] $x(t) = c_1 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} -6 \\ 1 \\ 5 \end{bmatrix} e^{-t} + c_3 \begin{bmatrix} -4 \\ 1 \\ 4 \end{bmatrix} e^{t}$

原点是鞍点，$c_3 = 0$ 的解被吸引到原点，$c_1 = c_2 = 0$ 的解被排斥.

17. [M]（复数解）：

$$c_1 \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} e^{t} + c_2 \begin{bmatrix} 23-34i \\ -9+14i \\ 3 \end{bmatrix} e^{(5+2i)t} + c_3 \begin{bmatrix} 23+34i \\ -9-14i \\ 3 \end{bmatrix} e^{(5-2i)t}$$

（实数解）：

$$c_1 \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} e^{t} + c_2 \begin{bmatrix} 23\cos 2t + 34\sin 2t \\ -9\cos 2t - 14\sin 2t \\ 3\cos 2t \end{bmatrix} e^{5t} +$$

$$c_3 \begin{bmatrix} 23\sin 2t - 34\cos 2t \\ -9\sin 2t + 14\cos 2t \\ 3\sin 2t \end{bmatrix} e^{5t}$$

原点是排斥子，轨迹是螺旋外出，远离原点.

19. [M] $A = \begin{bmatrix} -2 & 3/4 \\ 1 & -1 \end{bmatrix}$,

$$\begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} = \frac{5}{2}\begin{bmatrix} 1 \\ 2 \end{bmatrix}e^{-0.5t} - \frac{1}{2}\begin{bmatrix} -3 \\ 2 \end{bmatrix}e^{-2.5t}$$

21. [M] $A = \begin{bmatrix} -1 & -8 \\ 5 & -5 \end{bmatrix}$, $\begin{bmatrix} i_L(t) \\ v_C(t) \end{bmatrix} = \begin{bmatrix} -20\sin 6t \\ 15\cos 6t - 5\sin 6t \end{bmatrix}e^{-3t}$

5.8

1. 特征向量：$x_4 = \begin{bmatrix} 1 \\ 0.3326 \end{bmatrix}$，或 $Ax_4 = \begin{bmatrix} 4.9978 \\ 1.6652 \end{bmatrix}$；$\lambda \approx 4.9978$

3. 特征向量：$x_4 = \begin{bmatrix} 0.5188 \\ 1 \end{bmatrix}$，或 $Ax_4 = \begin{bmatrix} 0.4594 \\ 0.9075 \end{bmatrix}$；$\lambda \approx 0.9075$

5. $x = \begin{bmatrix} -0.7999 \\ 1 \end{bmatrix}, Ax = \begin{bmatrix} 4.0015 \\ -5.0020 \end{bmatrix}$;估计的 $\lambda = -5.0020$

7. [M] $x_k: \begin{bmatrix} 0.75 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0.9565 \end{bmatrix}, \begin{bmatrix} 0.9932 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0.9990 \end{bmatrix}, \begin{bmatrix} 0.9998 \\ 1 \end{bmatrix}$

   μ_k: 11.5, 12.78, 12.96, 12.9948, 12.9990

9. [M] $\mu_5 = 8.4233, \mu_6 = 8.4246$; 实际值：8.42443(精确到小数点后 5 位).

11. $\mu_k: 5.8000, 5.9655, 5.9942, 5.9990\ (k = 1, 2, 3, 4)$;
    $R(x_k): 5.9655, 5.9990, 5.99997, 5.9999993$

13. 是，但序列也许收敛非常慢.

15. 提示：$Ax - \alpha x = (A - \alpha I)x$，且用事实，当 α 不是 A 的特征值时，$(A - \alpha I)$ 是可逆的.

17. [M] $v_0 = 3.3384, v_1 = 3.32119$ (用四舍五入精确到 4 位有效数字)，$v_2 = 3.3212209$，实际值：3.3212201(精确到小数点后 7 位).

19. [M]a. $\mu_6 = 30.2887 = \mu_7$ 到四位小数，若到 6 位，最大特征值是 30.288685，特征向量为(0.957629，0.688937，1，0.943782).

    b. 逆幂法(取 $\alpha = 0$)得到 $\mu_1^{-1} = 0.010141$，$\mu_2^{-1} = 0.010150$，若到 7 位数字，最小特征值是 0.0101500，且特征向量是(−0.603972，1，−0.251135，0.148953). 收敛速度很快的原因是第二小的特征值接近 0.85.

21. a. 如果 A 的特征值在数量上都小于 1 且 $x \neq 0$，那么 $A^k x$ 对充分大的 k 趋于一个特征向量.

    b. 如果主特征值是 1，且 x 具有与对应特征向量方向一致的分量，那么 $\{A^k x\}$ 将收敛于那个特征向量的倍数.

    c. 如果 A 的特征值在数量上都大于 1，且 x 不是一个特征向量，那么从 $A^k x$ 到最近特征向量的距离在 $k \to \infty$ 时会增大.

第 5 章补充习题

1 . a. T　b. F　c. T　d. F　e. T
f. T　g. F　h. T　i. F　j. T
k. F　l. F　m. F　n. T　o. F
p. T　q. F　r. T　s. F　t. T
u. T　v. T　w. F　x. T

3. a. 假设 $Ax = \lambda x$，其中 $x \neq 0$，则 $(5I - A)x = 5x - Ax = 5x - \lambda x = (5 - \lambda)x$，特征值是 $5 - \lambda$.

   b. $(5I - 3A + A^2)x = 5x - 3Ax + A(Ax) = 5x - 3\lambda x + \lambda^2 x = (5 - 3\lambda + \lambda^2)x$，特征值为 $5 - 3\lambda + \lambda^2$.

5. 假设 $Ax = \lambda x$，$x \neq 0$，那么

$$\begin{aligned} p(A)x &= (c_0 I + c_1 A + c_2 A^2 + \cdots + c_n A^n)x \\ &= c_0 x + c_1(Ax) + c_2 A^2 x + \cdots + c_n A^n x \\ &= c_0 x + c_1 \lambda x + c_2 \lambda^2 x + \cdots + c_n \lambda^n x = p(\lambda)x \end{aligned}$$

   所以，$p(\lambda)$ 是 $p(A)$ 的一个特征值.

7. 如果 $A = PDP^{-1}$，那么 $p(A) = Pp(D)P^{-1}$，像习题 6 中所证明，如果 D 中 (j, j) 位置的元素为 λ，那么 D^k 中 (j, j) 位置为 λ^k.所以，$p(D)$ 中 (j, j) 位置的元素为 $p(\lambda)$. 如果 p 是矩阵 A 的特征多项式，那么对 D 的对角线每一个元素有 $p(\lambda) = 0$，因为 D 中的这些元素是 A 的特征值，于是 $p(D)$ 是零矩阵，从而 $p(A) = P \cdot 0 \cdot P^{-1} = 0$.

9. 如果 $I - A$ 不可逆，那么方程 $(I - A)x = 0$ 将会有一个非平凡解，从而 $x - Ax = 0$ 且 $Ax = 1 \cdot x$，这表明 A 具有特征值 1.如果所有特征值小于 1，这是不可能出现，所以 $I - A$ 一定可逆.

11. a. 取 H 中的 x，那么对一些数量 c 有 $x = cu$，所以 $Ax = A(cu) = c(Au) = c(\lambda u) = (c\lambda)u$，这说

明 $A\boldsymbol{x}$ 在 H 中.

b. 令 $\boldsymbol{x}$ 是 K 中的非零向量，由于 K 是一维的，K 是 $\boldsymbol{x}$ 的所有倍数形成的集合，如果 K 在 A 之下不变，那么 $A\boldsymbol{x}$ 属于 K，因此 $A\boldsymbol{x}$ 是 $\boldsymbol{x}$ 的倍数，那么 $\boldsymbol{x}$ 是 A 的一个特征向量.

13. 1，3，7

15. 将第 3 章补充习题 16 的行列式公式中的 a 用 $a-\lambda$ 替换：

$\det(A-\lambda I)=(a-b-\lambda)^{n-1}[a-\lambda+(n-1)b]$

这个行列式为零当且仅当 $a-b-\lambda=0$ 或 $a-\lambda+(n-1)b=0$. 于是 λ 是 A 的特征值当且仅当 $\lambda=a-b$ 或 $\lambda=a+(n-1)b$. 从上面计算 $\det(A-\lambda I)$ 的公式得到特征值 $a-b$ 的重数是 $n-1$，特征值 $a+(n-1)b$ 的重数是 1.

17. $\det(A-\lambda I)=(a_{11}-\lambda)(a_{22}-\lambda)-a_{12}a_{21}=$

$\lambda^2-(a_{11}+a_{22})\lambda+(a_{11}a_{22}-a_{12}a_{21})$

$=\lambda^2-(\operatorname{tr}A)\lambda+\det A$

利用二次求根公式求得特征方程的解为：

$$\lambda=\frac{\operatorname{tr}A\pm\sqrt{(\operatorname{tr}A)^2-4\det A}}{2}$$

特征值都是实数当且仅当判别式是非负的，即 $(\operatorname{tr}A)^2-4\det A\geqslant 0$. 这个不等式简化为 $(\operatorname{tr}A)^2\geqslant 4\det A$ 和 $\left(\frac{\operatorname{tr}A}{2}\right)^2\geqslant\det A$.

19. $C_p=\begin{bmatrix}0&1\\-6&5\end{bmatrix}$；$\det(C_p-\lambda I)=6-5\lambda+\lambda^2=p(\lambda)$

21. 如果 p 是一个次数为 2 的多项式，那么类似习题 19 中的一个计算表明：C_p 的特征多项式是 $p(\lambda)=(-1)^2p(\lambda)$，所以，对 $n=2$ 结果是真的，假若结果对 $n=k,k\geqslant 2$ 是真的，且考虑一个次数为 $(k+1)$ 的多项式 p. $\det(C_p-\lambda I)$ 按第一列的余因子展开，$C_p-\lambda I$ 的行列式等于

$$(-\lambda)\det\begin{bmatrix}-\lambda&1&\cdots&0\\\vdots&&&\vdots\\0&&&1\\-a_1&-a_2&\cdots&-a_k-\lambda\end{bmatrix}+(-1)^{k+1}a_0$$

显示的 $k\times k$ 矩阵是 $C_q-\lambda I$，此处 $q(t)=a_1+a_2t+\cdots+a_kt^{k-1}+t^k$. 由归纳假设，行列式 $C_p-\lambda I$ 是 $(-1)^kq(\lambda)$，行列式

$$\begin{aligned}\det(C_p-\lambda I)&=(-1)^{k+1}a_0+(-\lambda)\ (-1)^kq(\lambda)\\&=(-1)^{k+1}[a_0+\lambda(a_1+\cdots+a_k\lambda^{k-1}+\lambda^k)]\\&=(-1)^{k+1}p(\lambda)\end{aligned}$$

所以，当 $n=k$ 结论成立时，公式对 $n=k+1$ 成立. 由归纳法原理，关于 $\det(C_p-\lambda I)$ 的公式对所有 $n\geqslant 2$ 成立.

23. 由习题 22，范德蒙德矩阵 V 的列是 C_p 的特征向量，对应的特征值分别是 $\lambda_1,\lambda_2,\lambda_3$（特征多项式 p 的根），由于这些特征值是不同的，由 5.1 节定理 2，对应特征向量形成一个线性无关集. 于是 V 有线性无关的列，且由可逆矩阵定理可知 V 是可逆的. 最后，由于 V 的列是 C_p 的特征向量，对角化定理（5.3 节定理5）表明 $V^{-1}C_pV$ 是对角形.

25. [M]如果你的矩阵程序用迭代方法而不是符号计算法计算特征值和特征向量，你可能遇到一些困难. 你会发现 $AP-PD$ 有非常小的元素且 PDP^{-1} 接近于 A.（这在几年前是真的，但情况会随矩阵程序不断改进而改变.）如果你从程序的特征向量构造 P，注意检查 P 的条件数. 这也许告诉你不可以真正求得 3 个线性无关的特征向量. A 的特征值是小实数，如果你的矩阵程序使用无穷精确的符号计算，你会发现 A 是不可对角化.

第 6 章

6.1

1. 5，8，$\frac{8}{5}$　　3. $\begin{bmatrix}3/35\\-1/35\\-1/7\end{bmatrix}$　　5. $\begin{bmatrix}8/13\\12/13\end{bmatrix}$

7. $\sqrt{35}$　　9. $\begin{bmatrix}-0.6\\0.8\end{bmatrix}$　　11. $\begin{bmatrix}7/\sqrt{69}\\2/\sqrt{69}\\4/\sqrt{69}\end{bmatrix}$

13. $5\sqrt{5}$　　15. 不正交

17. 正交

19. 写出答案之后参考“学习指导”.

21. 提示：用 2.1 节的定理 3 和 2.

23. $u\cdot v=0$, $\|u\|^2=30$, $\|v\|^2=101$, $\|u+v\|^2$
$=(-5)^2+(-9)^2+5^2=131=30+101$.

25. 所有$\begin{bmatrix}-b\\a\end{bmatrix}$的倍数的集合(当 $v\neq 0$).

27. 提示：用正交性的定义.

29. 提示：考虑一个属于 W 的典型向量 $w=c_1v_1+\cdots+c_pv_p$

31. 提示：如果 x 属于 $W^\perp$，那么 x 与 W 中任一向量正交.

33. [M]说明你的猜测且给出代数验证.

6.2

1. 不正交　　3. 不正交　　5. 正交

7. 证明 $u_1\cdot u_2=0$，注意到定理 4 且观察到 $\mathbb{R}^2$ 中两个线性无关向量构成一个基，那么得到：
$$x=\frac{39}{13}\begin{bmatrix}2\\-3\end{bmatrix}+\frac{26}{52}\begin{bmatrix}6\\4\end{bmatrix}=3\begin{bmatrix}2\\-3\end{bmatrix}+\frac{1}{2}\begin{bmatrix}6\\4\end{bmatrix}$$

9. 证明 $u_1\cdot u_2=0$, $u_1\cdot u_3=0$ 且 $u_2\cdot u_3=0$，注意到定理 4 且观察到 $\mathbb{R}^3$ 中线性无关向量构成一个基，那么得到 $x=\frac{5}{2}u_1-\frac{27}{18}u_2+\frac{18}{9}u_3=\frac{5}{2}u_1-\frac{3}{2}u_2+2u_3$.

11. $\begin{bmatrix}-2\\1\end{bmatrix}$　　13. $y=\begin{bmatrix}-4/5\\7/5\end{bmatrix}+\begin{bmatrix}14/5\\8/5\end{bmatrix}$

15. $y-\hat{y}=\begin{bmatrix}0.6\\-0.8\end{bmatrix}$，距离是 1.

17. $\begin{bmatrix}1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{bmatrix}$, $\begin{bmatrix}-1/\sqrt{2}\\0\\1/\sqrt{2}\end{bmatrix}$

19. 标准正交　　21. 标准正交

23. 见“学习指导”.

25. 提示：$\|Ux\|^2=(Ux)^{\mathrm{T}}(Ux)$，同样，部分(a)和(c)可从部(b)得到.

27. 提示：你需要 2 个定理，其中一个仅应用于方阵.

29. 提示：如果你有一个候补逆，你可以检查这个候补是否可行.

31. 假若 $\hat{y}=\frac{y\cdot u}{u\cdot u}u$，对 $c\neq 0$ 用 cu 代替 u；那么
$$\frac{y\cdot(cu)}{(cu)\cdot(cu)}(cu)=\frac{c(y\cdot u)}{c^2u\cdot u}(c)u=\hat{y}$$

33. 设 $L=\mathrm{Span}\{u\}$，其中 u 非零，记 $T(x)=\mathrm{proj}_L x$，根据定义，$T(x)=\frac{x\cdot u}{u\cdot u}u=(x\cdot u)(u\cdot u)^{-1}u$
对 $\mathbb{R}^n$ 中的 x 和 y，任意数 c 和 d，内积的性质（定理 1）说明
$$\begin{aligned}T(cx+dy)&=[(cx+dy)\cdot u](u\cdot u)^{-1}u\\&=[c(x\cdot u)+d(y\cdot u)](u\cdot u)^{-1}u\\&=c(x\cdot u)(u\cdot u)^{-1}u+d(y\cdot u)(u\cdot u)^{-1}u\\&=cT(x)+dT(y)\end{aligned}$$
因此 T 是线性的.

6.3

1. $x=-\frac{8}{9}u_1-\frac{2}{9}u_2+\frac{2}{3}u_3+2u_4$；$x=\begin{bmatrix}0\\-2\\4\\-2\end{bmatrix}+\begin{bmatrix}10\\-6\\-2\\2\end{bmatrix}$

3. $\begin{bmatrix}-1\\4\\0\end{bmatrix}$　　5. $\begin{bmatrix}-1\\2\\6\end{bmatrix}=y$

7. $y=\begin{bmatrix}10/3\\2/3\\8/3\end{bmatrix}+\begin{bmatrix}-7/3\\7/3\\7/3\end{bmatrix}$　　9. $y=\begin{bmatrix}2\\4\\0\\0\end{bmatrix}+\begin{bmatrix}2\\-1\\3\\-1\end{bmatrix}$

11. $\begin{bmatrix}3\\-1\\1\\-1\end{bmatrix}$　　13. $\begin{bmatrix}-1\\-3\\-2\\3\end{bmatrix}$　　15. $\sqrt{40}$

17. a. $U^{\mathrm{T}}U=\begin{bmatrix}1&0\\0&1\end{bmatrix}$, $UU^{\mathrm{T}}=\begin{bmatrix}8/9&-2/9&2/9\\-2/9&5/9&4/9\\2/9&4/9&5/9\end{bmatrix}$

b. $\mathrm{proj}_W y=6u_1+3u_2=\begin{bmatrix}2\\4\\5\end{bmatrix}$，且 $(UU^{\mathrm{T}})y=\begin{bmatrix}2\\4\\5\end{bmatrix}$

19. $\begin{bmatrix}0\\2/5\\1/5\end{bmatrix}$的任何倍数，如$\begin{bmatrix}0\\2\\1\end{bmatrix}$.

21. 在查阅“学习指导”之前，写出答案.

23. 提示：用定理3和正交分解定理，对唯一性，假若 $A\boldsymbol{p}=\boldsymbol{b}$ 和 $A\boldsymbol{p}_1=\boldsymbol{b}$，考虑方程 $\boldsymbol{p}=\boldsymbol{p}_1+(\boldsymbol{p}-\boldsymbol{p}_1)$ 和 $\boldsymbol{p}=\boldsymbol{p}+\boldsymbol{0}$.

6.4

1. $\begin{bmatrix}3\\0\\-1\end{bmatrix},\begin{bmatrix}-1\\5\\-3\end{bmatrix}$ 3. $\begin{bmatrix}2\\-5\\1\end{bmatrix},\begin{bmatrix}3\\3/2\\3/2\end{bmatrix}$

5. $\begin{bmatrix}1\\-4\\0\\1\end{bmatrix},\begin{bmatrix}5\\1\\-4\\-1\end{bmatrix}$ 7. $\begin{bmatrix}2/\sqrt{30}\\-5/\sqrt{30}\\1/\sqrt{30}\end{bmatrix},\begin{bmatrix}2/\sqrt{6}\\1/\sqrt{6}\\1/\sqrt{6}\end{bmatrix}$

9. $\begin{bmatrix}3\\1\\-1\\3\end{bmatrix},\begin{bmatrix}1\\3\\3\\-1\end{bmatrix},\begin{bmatrix}-3\\1\\1\\3\end{bmatrix}$ 11. $\begin{bmatrix}1\\-1\\-1\\1\\1\end{bmatrix},\begin{bmatrix}3\\0\\3\\-3\\3\end{bmatrix},\begin{bmatrix}2\\0\\2\\2\\-2\end{bmatrix}$

13. $R=\begin{bmatrix}6&12\\0&6\end{bmatrix}$

15. $Q=\begin{bmatrix}1/\sqrt{5}&1/2&1/2\\-1/\sqrt{5}&0&0\\-1/\sqrt{5}&1/2&1/2\\1/\sqrt{5}&-1/2&1/2\\1/\sqrt{5}&1/2&-1/2\end{bmatrix},R=\begin{bmatrix}\sqrt{5}&-\sqrt{5}&4\sqrt{5}\\0&6&-2\\0&0&4\end{bmatrix}$

17. 见“学习指导”.

19. 假若 $\boldsymbol{x}$ 满足 $R\boldsymbol{x}=\boldsymbol{0}$，那么 $QR\boldsymbol{x}=Q\boldsymbol{0}=\boldsymbol{0}$，且 $A\boldsymbol{x}=\boldsymbol{0}$，由于 A 的列是线性无关，$\boldsymbol{x}$ 必须是零. 这个事实，反过来，说明 R 的列是线性无关，因为 R 是方阵，由可逆矩阵定理可知它是可逆的.

21. 记 Q 的列为 $\boldsymbol{q}_1,\cdots,\boldsymbol{q}_n$，注意 $n\leqslant m$，因为 A 是 $m\times n$ 且有线性无关的列，利用事实，Q 的列可以扩充为 $\mathbb{R}^m$ 的一个标准正交基，如 $\{\boldsymbol{q}_1,\cdots,\boldsymbol{q}_m\}$（“学习指导”描述了一个方法)，取 $Q_0=[\boldsymbol{q}_{n+1}\ \cdots\ \boldsymbol{q}_m]$ 那么 $Q_1=[Q\ \ Q_0]$，利用矩阵分解乘积 $Q_1\begin{bmatrix}R\\0\end{bmatrix}=QR=A$.

23. 提示：将 R 做为 2×2 分块矩阵.

25. [M] R 的对角元素是 20，6，10.3923 和 7.0711，精确到4位小数.

6.5

1. a. $\begin{bmatrix}6&-11\\-11&22\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}-4\\11\end{bmatrix}$ b. $\hat{\boldsymbol{x}}=\begin{bmatrix}3\\2\end{bmatrix}$

3. a. $\begin{bmatrix}6&6\\6&42\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}6\\-6\end{bmatrix}$ b. $\hat{\boldsymbol{x}}=\begin{bmatrix}4/3\\-1/3\end{bmatrix}$

5. $\hat{\boldsymbol{x}}=\begin{bmatrix}5\\-3\\0\end{bmatrix}+x_3\begin{bmatrix}-1\\1\\1\end{bmatrix}$ 7. $2\sqrt{5}$

9. a. $\hat{\boldsymbol{b}}=\begin{bmatrix}1\\1\\0\end{bmatrix}$ b. $\hat{\boldsymbol{x}}=\begin{bmatrix}2/7\\1/7\end{bmatrix}$

11. a. $\hat{\boldsymbol{b}}=\begin{bmatrix}3\\1\\4\\-1\end{bmatrix}$ b. $\hat{\boldsymbol{x}}=\begin{bmatrix}2/3\\0\\1/3\end{bmatrix}$

13. $A\boldsymbol{u}=\begin{bmatrix}11\\-11\\11\end{bmatrix}$, $A\boldsymbol{v}=\begin{bmatrix}7\\-12\\7\end{bmatrix}$, $\boldsymbol{b}-A\boldsymbol{u}=\begin{bmatrix}0\\2\\-6\end{bmatrix}$, $\boldsymbol{b}-A\boldsymbol{v}=\begin{bmatrix}4\\3\\-2\end{bmatrix}$，否，$\boldsymbol{u}$ 不可能是 $A\boldsymbol{x}=\boldsymbol{b}$ 的一个最小二乘解. 为什么？

15. $\hat{\boldsymbol{x}}=\begin{bmatrix}4\\-1\end{bmatrix}$ 17. 见“学习指导”.

19. a. 如果 $A\boldsymbol{x}=\boldsymbol{0}$，那么 $A^TA\boldsymbol{x}=A^T\boldsymbol{0}=\boldsymbol{0}$，这说明 Nul A 包含在 Nul A^TA 中.

 b. 如果 $A^TA\boldsymbol{x}=\boldsymbol{0}$，那么 $\boldsymbol{x}^TA^TA\boldsymbol{x}=\boldsymbol{x}^T\boldsymbol{0}=0$，所以，$(A\boldsymbol{x})^T(A\boldsymbol{x})=0$（这说明 $\|A\boldsymbol{x}\|^2=0$），因此 $A\boldsymbol{x}=\boldsymbol{0}$，这说明 Nul A^TA 包含在 Nul A 中.

21. 提示：对(a)，利用第2章的一个重要定理.

23. 由定理 14，$\hat{\boldsymbol{b}}=A\hat{\boldsymbol{x}}=A(A^TA)^{-1}A^T\boldsymbol{b}$，矩阵 $A(A^TA)^{-1}A^T$ 在统计中经常出现，通常称为帽子矩阵.

25. 标准方程是 $\begin{bmatrix}2&2\\2&2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}6\\6\end{bmatrix}$，它的解是 (x,y) 的集合使得 $x+y=3$. 解对应位于直线

$x+y=2$ 和 $x+y=4$ 中间的点集.

6.6

1. $y=0.9+0.4x$　　3. $y=1.1+1.3x$

5. 如果两个点集有不同的 x 坐标，那么设计矩阵 X 的两列不可能是相互是倍数，因而它们线性无关.由 6.5 节的定理 14，标准方程有唯一解.

7. a. $\boldsymbol{y}=X\boldsymbol{\beta}+\boldsymbol{\varepsilon}$，其中 $\boldsymbol{y}=\begin{bmatrix}1.8\\2.7\\3.4\\3.8\\3.9\end{bmatrix}$, $X=\begin{bmatrix}1&1\\2&4\\3&9\\4&16\\5&25\end{bmatrix}$,

$$\boldsymbol{\beta}=\begin{bmatrix}\beta_1\\\beta_2\end{bmatrix},\ \boldsymbol{\varepsilon}=\begin{bmatrix}\varepsilon_1\\\varepsilon_2\\\varepsilon_3\\\varepsilon_4\\\varepsilon_5\end{bmatrix}$$

b. [M] $y=1.76x-0.20x^2$

9. $\boldsymbol{y}=X\boldsymbol{\beta}+\boldsymbol{\varepsilon}$，其中 $\boldsymbol{y}=\begin{bmatrix}7.9\\5.4\\-0.9\end{bmatrix}$, $X=\begin{bmatrix}\cos 1&\sin 1\\\cos 2&\sin 2\\\cos 3&\sin 3\end{bmatrix}$,

$$\boldsymbol{\beta}=\begin{bmatrix}A\\B\end{bmatrix},\ \boldsymbol{\varepsilon}=\begin{bmatrix}\varepsilon_1\\\varepsilon_2\\\varepsilon_3\end{bmatrix}$$

11. [M] $\beta=1.45$ 和 $e=0.811$；轨迹一个椭圆，方程 $r=\beta/(1-e\cdot\cos\ \vartheta)$，当 $\vartheta=4.6$ 时，$r=1.33$.

13. [M] a. $y=-0.8558+4.7025t+5.5554t^2-0.0274t^3$

b. 速度函数是

$v(t)=4.7025+11.1108t-0.0822t^2$

$v(4.5)=53.0$ 英尺/秒

15. 提示：写出在方程(1)中的 X 和 $\boldsymbol{y}$，且计算 $X^{\mathrm{T}}X$ 和 $X^{\mathrm{T}}\boldsymbol{y}$.

17. a. 数据 $\boldsymbol{x}$ 的平均 $\bar{x}=5.5$，数据的平均偏差形式是 $(-3.5,1),(-0.5,2),(1.5,3),(2.5,3)$. X 的列是正交的,因为第 2 列中元素之和为零.

b. $\begin{bmatrix}4&0\\0&21\end{bmatrix}\begin{bmatrix}\beta_0\\\beta_1\end{bmatrix}=\begin{bmatrix}9\\7.5\end{bmatrix}$, $y=\dfrac{9}{4}+\dfrac{5}{14}x^*=\dfrac{9}{4}+\dfrac{5}{14}(x-5.5)$

19. 提示：方程有一个好的几何解释.

6.7

1. a. 3，$\sqrt{105}$，225　　b. 所有 $\begin{bmatrix}1\\4\end{bmatrix}$ 的倍数

3. 28　　5. $5\sqrt{2}$，$3\sqrt{3}$　　7. $\dfrac{56}{25}+\dfrac{14}{25}t$

9. a. 常数多项式，$p(t)=5$

b. t^2-5 是正交于 p_0 和 p_1，值：$(4,-4,-4,4)$；

答案：$q(t)=\dfrac{1}{4}(t^2-5)$

11. $\dfrac{17}{5}t$

13. 验证 4 个公理中的每一个，例如

1. $\langle\boldsymbol{u},\boldsymbol{v}\rangle=(A\boldsymbol{u})\cdot(A\boldsymbol{v})$　定义

$=(A\boldsymbol{v})\cdot(A\boldsymbol{u})$　点积的性质

$=\langle\boldsymbol{v},\boldsymbol{u}\rangle$　定义

15. $\langle\boldsymbol{u},c\boldsymbol{v}\rangle=\langle c\boldsymbol{v},\boldsymbol{u}\rangle$　公理1

$=c\langle\boldsymbol{v},\boldsymbol{u}\rangle$　公理3

$=c\langle\boldsymbol{u},\boldsymbol{v}\rangle$　公理1

17. 提示：计算 4 乘右手边.

19. $\langle\boldsymbol{u},\boldsymbol{v}\rangle=\sqrt{a}\sqrt{b}+\sqrt{b}\sqrt{a}=2\sqrt{ab}$，

$\|\boldsymbol{u}\|^2=(\sqrt{a})^2+(\sqrt{b})^2=a+b$，因为 a,b 非负，所以 $\|\boldsymbol{u}\|=\sqrt{a+b}$，类似，$\|\boldsymbol{v}\|=\sqrt{b+a}$，由柯西-施瓦茨不等式，$2\sqrt{ab}\leqslant\sqrt{a+b}\sqrt{b+a}=a+b$，因而 $\sqrt{ab}\leqslant\dfrac{a+b}{2}$.

21. 0　　23. $2/\sqrt{5}$　　25. 1，t，$3t^2-1$

27. [M] 新的正交多项式是 $-17t+5t^3$ 和 $72-155t^2+35t^4$ 的倍数，重新度量这些多项式使得它们在-2，-1，0，1 和 2 的值是小整数.

6.8

1. $y=2+\dfrac{3}{2}t$

3. $p(t)=4p_0-0.1p_1-0.5p_2+0.2p_3$

$$=4-0.1t-0.5(t^2-2)+0.2\left(\frac{5}{6}t^3-\frac{17}{6}t\right)$$

（这个多项式精确地拟合了数据.）

5. 利用等式

$$\sin mt \ \sin nt = \frac{1}{2}[\cos(mt-nt)-\cos(mt+nt)]$$

7. 利用等式 $\cos^2 kt = \dfrac{1+\cos 2kt}{2}$.

9. $\pi + 2\sin t + \sin 2t + \frac{2}{3}\sin 3t$（提示：利用例 4 的结果可节约时间）.

11. $\frac{1}{2} - \frac{1}{2}\cos 2t$（为什么？）

13. 提示：选取 $C[0,2\pi]$ 中的函数 f 和 g，且固定一个整数 $m \geqslant 0$. 写出 $f+g$ 的包含 cos mt 的傅里叶系数且写出包含 sin mt 的傅里叶系数 $(m>0)$.

15. [M]立方曲线是下列函数的图像 $g(t) = -0.2685 + 3.6095t + 5.8576t^2 - 0.0477t^3$. 在 $t=4.5$ 秒速度是 $g'(4.5)=53.4$ 英尺/秒.这比 6.6 节习题 13 中估计得到的结果快 0.7%.

第 6 章补充习题

1. a. F　b. T　c. T　d. F　e. F
   f. T　g. T　h. T　i. F　j. T
   k. T　l. F　m. T　n. F　o. F
   p. T　q. T　r. F　s. F

2. 提示：若 $\{v_1, v_2\}$ 是正交集合且 $x = c_1v_1 + c_2v_2$，则向量 c_1v_1, c_2v_2 是正交的且
$$\|x\|^2 = \|c_1v_1 + c_2v_2\|^2 = \|c_1v_1\|^2 + \|c_2v_2\|^2$$
$$= (|c_1|\|v_1\|)^2 + (|c_2|\|v_2\|)^2 = |c_1|^2 + |c_2|^2$$
故对于 $p=2$ 状态方程成立. 假设等式对 $p=k$ $(k\geqslant 2)$ 成立，令 $\{v_1,\cdots,v_{k+1}\}$ 是标准正交集，考虑 $x = c_1v_1+\cdots+c_kv_k+c_{k+1}v_{k+1} = u_k + c_{k+1}v_{k+1}$，这里 $u_k = c_1v_1+\cdots+c_kv_k$.

3. 给定 x 和 $\mathbb{R}^n$ 中一个标准正交基 $\{v_1,\cdots,v_p\}$，设 $\hat{x}$ 是 x 在由 $v_1,\cdots,v_p$ 生成的子空间上的正交投影，由 6.3 节的定理 10，$\hat{x} = (x\cdot v_1)v_1 + \cdots + (x\cdot v_p)v_p$. 由习题 2，$\|\hat{x}\|^2 = |x\cdot v_1|^2 + \cdots + |x\cdot v_p|^2$. 贝塞尔不等式可从 $\|\hat{x}\|^2 \leqslant \|x\|^2$ 得到，该不等式在 6.7 节柯西-施瓦茨不等式的证明前面给出.

5. 假若对任意 x, y 属于 $\mathbb{R}^n$ 有 $(Ux)\cdot(Uy) = x\cdot y$，且设 $e_1,\cdots,e_n$ 是 $\mathbb{R}^n$ 的标准基，对 $j=1,\cdots,n$，Ue_j 是 U 的第 j 列，由于 $\|Ue_j\|^2 = (Ue_j)\cdot(Ue_j) = e_j\cdot e_j = 1$，$U$ 的列是单位向量，由于对 $j\neq k$，$(Ue_j)\cdot(Ue_k) = e_j\cdot e_k = 0$，列是两两正交.

7. 提示：计算 Q^TQ，利用事实 $(uu^T)^T = u^{TT}u^T = uu^T$.

9. 设 $W = \text{Span}\{u, v\}$. 给定 $\mathbb{R}^n$ 中的 z，设 $\hat{z} = \text{proj}_W z$，那么 $\hat{z}$ 属于 Col A，此处 $A = [u \ \ v]$，例如，对 $\mathbb{R}^2$ 中的一些 $\hat{x}$，有 $\hat{z} = A\hat{x}$，所以 $\hat{x}$ 是 $Ax = z$ 的一个最小二乘解，标准方程可以解出 $\hat{x}$，那么 $\hat{z}$ 可通过计算 $A\hat{x}$ 得到.

11. 提示：设 $x = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, $b = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, $v = \begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix}$ 和 $A = \begin{bmatrix} v^T \\ v^T \\ v^T \end{bmatrix} = \begin{bmatrix} 1 & -2 & 5 \\ 1 & -2 & 5 \\ 1 & -2 & 5 \end{bmatrix}$. 给定的方程组是 $Ax = b$，且所有最小二乘解的集合与 $A^TAx = A^Tb$ 的解集相一致（6.5 节定理 13），这可利用事实 $(vv^T)x = v(x^Tx) = (v^Tx)v$，其中 v^Tx 是数.

13. a. Au 的行一列计算表明 A 的每一行是与 Nul A 中的每一个 u 正交的.所以 A 的每一行是属于 $(\text{Nul } A)^\perp$. 由于 $(\text{Nul } A)^\perp$ 是一个子空间，它必须包含 A 的行的所有线性组合；因此 $(\text{Nul } A)^\perp$ 包含 Row A.

    b. 如果 rank $A = r$，那么由秩定理 dim Nul $A = n - r$，由 6.3 节的习题 24(c)，有
    dim Nul $A + \dim(\text{Nul } A)^\perp = n$
    所以 $\dim(\text{Nul } A)^\perp$ 一定是 r，但由秩定理和(a)可知 Row A 是一个 $(\text{Nul } A)^\perp$ 的 r-维子空间，所以 Row A 必须和 $(\text{Nul } A)^\perp$ 一致.

    c. 在(b)中用 A^T 代替 A 且可得 Row A^T 和 $(\text{Nul } A^T)^\perp$ 一致，又由于 Row $A^T = $ Col A，这就证明了(c).

15. 如果 $A = URU^T$ 且 U 是正交的，那么 A 与 R 相似(因为 U 是可逆的且 $U^T = U^{-1}$)，且 A 与 R 具有相同特征值(5.2 节定理 4)；就是 n 个在 R 对

角线上的实数.

17. [M] $\frac{\|\Delta x\|}{\|x\|}=0.4618$, $\text{cond}(A)\times\frac{\|\Delta b\|}{\|b\|}=3363\times(1.548\times10^{-4})=0.5206$.
注意 $\|\Delta x\|/\|x\|$ 约等于 $\text{cond}(A)$ 乘 $\|\Delta b\|/\|b\|$.

19. [M] $\frac{\|\Delta x\|}{\|x\|}=7.178\times10^{-8}$, $\frac{\|\Delta b\|}{\|b\|}=2.832\times10^{-4}$，注意 x 的相对改变比 b 的相对改变小很多，事实上，由于

$$\text{cond}(A)\times\frac{\|\Delta b\|}{\|b\|}=23683\times(2.832\times10^{-4})=6.707$$

x 相对改变的理论上界是 6.707(到 4 位有效数字).这个习题表明，既使一个条件数很大，一个解的相对误差并不像你想象的那么大.

第 7 章

7.1

1. 对称　3. 不对称　5. 不对称

7. 正交，$\begin{bmatrix}0.6 & 0.8\\0.8 & -0.6\end{bmatrix}$　9. 不正交

11. 正交，$\begin{bmatrix}2/3 & 0 & 5/\sqrt{45}\\2/3 & 1/\sqrt{5} & -4/\sqrt{45}\\1/3 & -2/\sqrt{5} & -2/\sqrt{45}\end{bmatrix}$

13. $P=\begin{bmatrix}1/\sqrt{2} & -1/\sqrt{2}\\1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}$, $D=\begin{bmatrix}4 & 0\\0 & 2\end{bmatrix}$

15. $P=\begin{bmatrix}-4/\sqrt{17} & 1/\sqrt{17}\\1/\sqrt{17} & 4/\sqrt{17}\end{bmatrix}$, $D=\begin{bmatrix}17 & 0\\0 & 0\end{bmatrix}$

17. $P=\begin{bmatrix}1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2}\\1/\sqrt{3} & -2/\sqrt{6} & 0\\1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2}\end{bmatrix}$, $D=\begin{bmatrix}5 & 0 & 0\\0 & 2 & 0\\0 & 0 & -2\end{bmatrix}$

19. $P=\begin{bmatrix}-1/\sqrt{5} & 4/\sqrt{45} & -2/3\\2/\sqrt{5} & 2/\sqrt{45} & -1/3\\0 & 5/\sqrt{45} & 2/3\end{bmatrix}$, $D=\begin{bmatrix}7 & 0 & 0\\0 & 7 & 0\\0 & 0 & -2\end{bmatrix}$

21. $P=\begin{bmatrix}0.5 & -0.5 & -1/\sqrt{2} & 0\\0.5 & 0.5 & 0 & -1/\sqrt{2}\\0.5 & -0.5 & 1/\sqrt{2} & 0\\0.5 & 0.5 & 0 & 1/\sqrt{2}\end{bmatrix}$,
$D=\begin{bmatrix}9 & 0 & 0 & 0\\0 & 5 & 0 & 0\\0 & 0 & 1 & 0\\0 & 0 & 0 & 1\end{bmatrix}$

23. $P=\begin{bmatrix}1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6}\\1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6}\\1/\sqrt{3} & 0 & 2/\sqrt{6}\end{bmatrix}$, $D=\begin{bmatrix}5 & 0 & 0\\0 & 2 & 0\\0 & 0 & 2\end{bmatrix}$

25. 见“学习指导”.

27. $(B^{T}AB)^{T}=B^{T}A^{T}B^{TT}$　转置乘积的顺序是原乘积的相反顺序
$=B^{T}AB$　由于 A 对称
当 $A=I$ 时关于 $B^{T}B$ 的结果是一个特殊情况，$(BB^{T})^{T}=B^{TT}B^{T}=BB^{T}$，所以 BB^{T} 是对称的.

29. 提示：用 A 的一个正交对角化，或用定理 2.

31. 5.3 节的对角化定理说明 P 的列是(线性无关的) D 中对角形列出的 A 的特征值对应的特征向量，所以 P 正好有 k 列特征向量对应 λ，这 k 列形成特征空间的一个基.

33. $A=8u_1u_1^{T}+6u_2u_2^{T}+3u_3u_3^{T}$
$$=8\begin{bmatrix}1/2 & -1/2 & 0\\-1/2 & 1/2 & 0\\0 & 0 & 0\end{bmatrix}+6\begin{bmatrix}1/6 & 1/6 & -2/6\\1/6 & 1/6 & -2/6\\-2/6 & -2/6 & 4/6\end{bmatrix}+3\begin{bmatrix}1/3 & 1/3 & 1/3\\1/3 & 1/3 & 1/3\\1/3 & 1/3 & 1/3\end{bmatrix}$$

35. 提示：$(uu^{T})x=u(u^{T}x)=(u^{T}x)u$，由于 $u^{T}x$ 是一个数.

7.2

1. a. $5x_1^2+\frac{2}{3}x_1x_2+x_2^2$
b. 185　c. 16

3. a. $\begin{bmatrix}10 & -3\\-3 & -3\end{bmatrix}$　b. $\begin{bmatrix}5 & 3/2\\3/2 & 0\end{bmatrix}$

5. a. $\begin{bmatrix} 8 & -3 & 2 \\ -3 & 7 & -1 \\ 2 & -1 & -3 \end{bmatrix}$ b. $\begin{bmatrix} 0 & 2 & 3 \\ 2 & 0 & -4 \\ 3 & -4 & 0 \end{bmatrix}$

7. $\boldsymbol{x} = P\boldsymbol{y}$，此处 $P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$，$\boldsymbol{y}^{\mathrm{T}}D\boldsymbol{y} = 6y_1^2 - 4y_2^2$

在习题 9~14 中，其他答案（坐标变换和新的二次型）也有可能.

9. 正定，特征值是 7 和 2

坐标改变：$\boldsymbol{x} = P\boldsymbol{y}$，其中 $P = \frac{1}{\sqrt{5}}\begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix}$

新的二次型：$7y_1^2 + 2y_2^2$

11. 不定，特征值是 7 和–3

坐标变换：$\boldsymbol{x} = P\boldsymbol{y}$，其中 $P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

新的二次型：$7y_1^2 - 3y_2^2$

13. 半正定，特征值是 10 和 0

坐标变换：$\boldsymbol{x} = P\boldsymbol{y}$，其中 $P = \frac{1}{\sqrt{10}}\begin{bmatrix} 1 & 3 \\ -3 & 1 \end{bmatrix}$

新的二次型：$10y_1^2$

15. [M] 半负定：特征值是：0，–6，–8，–12，坐标变换：$\boldsymbol{x} = P\boldsymbol{y}$

$$P = \begin{bmatrix} 3/\sqrt{12} & 0 & -1/2 & 0 \\ 1/\sqrt{12} & -2/\sqrt{6} & 1/2 & 0 \\ 1/\sqrt{12} & 1/\sqrt{6} & 1/2 & -1/\sqrt{2} \\ 1/\sqrt{12} & 1/\sqrt{6} & 1/2 & 1/\sqrt{2} \end{bmatrix}$$

新的二次型：$-6y_2^2 - 8y_3^2 - 12y_4^2$

17. [M] 不定，特征值是 8.5，–6.5

坐标变换 $\boldsymbol{x} = P\boldsymbol{y}$：

$$P = \frac{1}{\sqrt{50}}\begin{bmatrix} 3 & -4 & 3 & 4 \\ 5 & 0 & -5 & 0 \\ 4 & 3 & 4 & -3 \\ 0 & 5 & 0 & 5 \end{bmatrix}$$

新的二次型：$8.5y_1^2 + 8.5y_2^2 - 6.5y_3^2 - 6.5y_4^2$

19. 8

21. 见“学习指导”.

23. 用两种方式写出特征多项式

$$\det(A - \lambda I) = \det\begin{bmatrix} a-\lambda & b \\ b & d-\lambda \end{bmatrix}$$

$$= \lambda^2 - (a+d)\lambda + ad - b^2$$

$$(\lambda - \lambda_1)(\lambda - \lambda_2) = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2$$

系数相等得到 $\lambda_1 + \lambda_2 = a + d$ 和 $\lambda_1\lambda_2 = ad - b^2 = \det A$.

25. 7.1 节的习题 27 表明 $B^{\mathrm{T}}B$ 是对称的，并且，$\boldsymbol{x}^{\mathrm{T}}B^{\mathrm{T}}B\boldsymbol{x} = (B\boldsymbol{x})^{\mathrm{T}}B\boldsymbol{x} = \|B\boldsymbol{x}\|^2 \geqslant 0$，所以二次型是半正定，并且我们说矩阵 $B^{\mathrm{T}}B$ 是半正定，提示：为证明当 B 是方阵且可逆时 $B^{\mathrm{T}}B$ 是正定，假若 $\boldsymbol{x}^{\mathrm{T}}B^{\mathrm{T}}B\boldsymbol{x} = 0$ 然后推出 $\boldsymbol{x} = \boldsymbol{0}$.

27. 提示：证明 $A + B$ 是对称的且二次型 $\boldsymbol{x}^{\mathrm{T}}(A+B)\boldsymbol{x}$ 是正定的.

7.3

1. $\boldsymbol{x} = P\boldsymbol{y}$，此处 $P = \begin{bmatrix} 1/3 & 2/3 & -2/3 \\ 2/3 & 1/3 & 2/3 \\ -2/3 & 2/3 & 1/3 \end{bmatrix}$

3. a. 9 b. $\pm\begin{bmatrix} 1/3 \\ 2/3 \\ -2/3 \end{bmatrix}$ c.6

5. a. 7 b. $\pm\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ c.3

7. $\pm\begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}$ 9. $5 + \sqrt{5}$ 11. 3

13. 提示：如果 $m = M$，对 $\boldsymbol{x}$ 取公式中的 $\alpha = 0$，也就是说，取 $\boldsymbol{x} = \boldsymbol{u}_n$ 且验证 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} = m$.如果 $m < M$ 且如果 t 是介于 m 和 M 中间的数，那么 $0 \leqslant t - m \leqslant M - m$ 且 $0 \leqslant (t-m)/(M-m) \leqslant 1$，所以，设 $\alpha = (t-m)/(M-m)$，解这个 α 的表达式得到 $t = (1-\alpha)m + \alpha M$，当 α 变化从 0 到 1，t 从 m 变化到 M，像习题中所陈述的那样构造 $\boldsymbol{x}$ 且验证它的性质.

15. [M] a.7.5 b. $\begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \end{bmatrix}$ c. –0.5

17. [M] a. –4 b. $\begin{bmatrix} -3/\sqrt{12} \\ 1/\sqrt{12} \\ 1/\sqrt{12} \\ 1/\sqrt{12} \end{bmatrix}$ c. –10

7.4

1. 3，1　　3. 3，2

习题 5~13 中的答案不是唯一的.

5. $\begin{bmatrix}-3 & 0\\ 0 & 0\end{bmatrix}=\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}3 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$

7. $\begin{bmatrix}1/\sqrt{5} & -2/\sqrt{5}\\ 2/\sqrt{5} & 1/\sqrt{5}\end{bmatrix}\begin{bmatrix}3 & 0\\ 0 & 2\end{bmatrix}\begin{bmatrix}2/\sqrt{5} & 1/\sqrt{5}\\ -1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}$

9. $\begin{bmatrix}1/\sqrt{2} & -1/\sqrt{2} & 0\\ 0 & 0 & 1\\ 1/\sqrt{2} & 1/\sqrt{2} & 0\end{bmatrix}\begin{bmatrix}3\sqrt{10} & 0\\ 0 & \sqrt{10}\\ 0 & 0\end{bmatrix}$

$\times\begin{bmatrix}2/\sqrt{5} & 1/\sqrt{5}\\ -1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}$

11. $\begin{bmatrix}-1/3 & 2/3 & 2/3\\ 2/3 & -1/3 & 2/3\\ 2/3 & 2/3 & -1/3\end{bmatrix}\begin{bmatrix}3\sqrt{10} & 0\\ 0 & 0\\ 0 & 0\end{bmatrix}$

$\times\begin{bmatrix}3/\sqrt{10} & -1/\sqrt{10}\\ 1/\sqrt{10} & 3/\sqrt{10}\end{bmatrix}$

13. $\begin{bmatrix}3 & 2 & 2\\ 2 & 3 & -2\end{bmatrix}=\begin{bmatrix}1/\sqrt{2} & -1/\sqrt{2}\\ 1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}\begin{bmatrix}5 & 0 & 0\\ 0 & 3 & 0\end{bmatrix}$

$\times\begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2} & 0\\ -1/\sqrt{18} & 1/\sqrt{18} & -4/\sqrt{18}\\ -2/3 & 2/3 & 1/3\end{bmatrix}$

15. a. rank $A=2$

b. Col A 的基：$\begin{bmatrix}0.40\\ 0.37\\ -0.84\end{bmatrix},\begin{bmatrix}-0.78\\ -0.33\\ -0.52\end{bmatrix}$

Nul A 的基：$\begin{bmatrix}0.58\\ -0.58\\ 0.58\end{bmatrix}$（记住 V^{T} 出现在 SVD 中.）

17. 令 $A=U\Sigma V^{\mathrm{T}}=U\Sigma V^{-1}$，由于 A 是方阵且可逆，rank $A=n$，且 Σ 中所有对角元素必须非零，所以 $A^{-1}=(U\Sigma V^{-1})^{-1}=V\Sigma^{-1}U^{-1}=V\Sigma^{-1}U^{\mathrm{T}}$.

19. 提示：由于 U 和 V 是正交的，

$A^{\mathrm{T}}A=(U\Sigma V^{\mathrm{T}})^{\mathrm{T}}U\Sigma V^{\mathrm{T}}=V\Sigma^{\mathrm{T}}U^{\mathrm{T}}U\Sigma V^{\mathrm{T}}$

$=V(\Sigma^{\mathrm{T}}\Sigma)V^{-1}$

这样 V 将 $A^{\mathrm{T}}A$ 对角化，这里 V 告诉你些什么呢？

21. 设 $A=U\Sigma V^{\mathrm{T}}$，矩阵 PU 是正交的，由于 P 和 U 都正交（见 6.2 节习题 29），所以方程 $PA=(PU)\Sigma V^{\mathrm{T}}$ 有奇异分解值需要的形式.由习题 19，Σ 中的对角元素是 PA 的奇异值.

23. 提示：利用 $(U\Sigma)V^{\mathrm{T}}$ 的列行展开.

25. 提示：考虑 T 的 SVD 的标准矩阵，如 $A=U\Sigma V^{\mathrm{T}}=U\Sigma V^{-1}$，设 $\mathcal{B}=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$ 和 $\mathcal{C}=\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_m\}$ 分别是 V 和 U 的列构成的基.像 5.4 节计算 T 相对于 $\mathcal{B}$ 和 $\mathcal{C}$ 的矩阵，为进行这些，你必须证明 $V^{-1}\boldsymbol{v}_j=\boldsymbol{e}_j$，$I_n$ 的第 j 列.

27. [M] $\begin{bmatrix}-0.57 & -0.65 & -0.42 & 0.27\\ 0.63 & -0.24 & -0.68 & -0.29\\ 0.07 & -0.63 & 0.53 & -0.56\\ -0.51 & -0.34 & -0.29 & -0.73\end{bmatrix}$

$\times\begin{bmatrix}16.46 & 0 & 0 & 0 & 0\\ 0 & 12.16 & 0 & 0 & 0\\ 0 & 0 & 4.87 & 0 & 0\\ 0 & 0 & 0 & 4.31 & 0\end{bmatrix}$

$\times\begin{bmatrix}-0.10 & 0.61 & -0.21 & -0.52 & 0.55\\ -0.39 & 0.29 & 0.84 & -0.14 & -0.19\\ -0.74 & -0.27 & -0.07 & 0.38 & 0.49\\ 0.41 & -0.50 & 0.45 & -0.23 & 0.58\\ -0.36 & -0.48 & -0.19 & -0.72 & -0.29\end{bmatrix}$

29. [M]25.934 3, 16.755 4, 11.291 7, 1.078 5, 0.003 779 3; $\sigma_1/\sigma_5=68\,622$.

7.5

1. $M=\begin{bmatrix}12\\ 10\end{bmatrix}$; $B=\begin{bmatrix}7 & 10 & -6 & -9 & -10 & 8\\ 2 & -4 & -1 & 5 & 3 & -5\end{bmatrix}$;

$S=\begin{bmatrix}86 & -27\\ -27 & 16\end{bmatrix}$

3. $\begin{bmatrix}0.95\\ -0.32\end{bmatrix}$ 对应 $\lambda=95.2$，$\begin{bmatrix}0.32\\ 0.95\end{bmatrix}$ 对应 $\lambda=6.8$

5. [M](0.130，0.874，0.468)，方差的 75.9%.

7. $y_1=0.95x_1-0.32x_2$；y_1 解释方差的 93.3%.

9. $c_1=1/3$，$c_2=2/3$，$c_3=2/3$，y 的方差是 9.

11. a. 如果 $\boldsymbol{w}$ 是 $\mathbb{R}^N$ 中的向量，其中每个位置都是 1，因为 $\boldsymbol{X}_k$ 是平均偏差形式，那么

$$[X_1 \ \cdots \ X_N]\, w = X_1 + \cdots + X_N = 0$$

于是

$$\begin{aligned}[Y_1 \ \cdots \ Y_N]w &= [P^{\mathrm{T}}X_1 \ \cdots \ P^{\mathrm{T}}X_N]w \text{（根据定义）}\\ &= P^{\mathrm{T}}[X_1 \ \cdots \ X_N]w = P^{\mathrm{T}}0 = 0\end{aligned}$$

也就是说，$Y_1 + \cdots + Y_N = 0$，所以 Y_k 是平均偏差形式.

b. 提示：由于 X_j 是平均偏差形式，X_j 的协方差矩阵是 $1/(N-1)[X_1 \cdots X_N][X_1 \cdots X_N]^{\mathrm{T}}$，利用(a)，计算 Y_j 的协方差矩阵.

13. 如果 $B = [\hat{X}_1 \cdots \hat{X}_N]$，那么

$$S = \frac{1}{N-1}BB^{\mathrm{T}} = \frac{1}{N-1}\left[\hat{X}_1 \cdots \hat{X}_n\right]\begin{bmatrix}\hat{X}_1^{\mathrm{T}}\\ \vdots \\ \hat{X}_N^{\mathrm{T}}\end{bmatrix}$$

$$= \frac{1}{N-1}\sum_{k=1}^{N}\hat{X}_k\hat{X}_k^{\mathrm{T}} = \frac{1}{N-1}\sum_{k=1}^{N}(X_k - M)(X_k - M)^{\mathrm{T}}$$

第 7 章补充习题

1. a. T　b. F　c. T　d. F　e. F　f. F　g. F　h. T　i. F　j. F　k. F　l. F　m. T　n. F　o. T　p. T　q. F

3. 如果 $rank\ A = r$，那么由秩定理 dim Nul $A = n - r$，所以，0 是重数为 $n-r$ 的特征值，因此，A 的谱分解中的 n 项，正好有 $(n-r)$ 个为零，其余的 r 项，(对应非零特征值)都是秩为 1 的矩阵，像谱分解中讨论的一样.

5. 如果对一些非零 λ 有 $Av = \lambda v$，那么 $v = \lambda^{-1}Av = A(\lambda^{-1}v)$，它证明 v 是 A 的列的一个线性组合.

7. 提示：如果 $A = R^{\mathrm{T}}R$，其中 R 是可逆的，由 7.2 节的习题 25，那么 A 是正定的.相反地，假若 A 是正定，那么 7.2 节的习题 26，对一些正定矩阵 B 有 $A = B^{\mathrm{T}}B$.解释为什么 B 有一个 QR 分解，且利用它产生 A 的一个乔累斯基分解.

9. 如果 A 是一个 $m \times n$ 且 x 属于 $\mathbb{R}^n$，那么 $x^{\mathrm{T}}A^{\mathrm{T}}Ax = (Ax)^{\mathrm{T}}(Ax) = \|Ax\|^2 \geqslant 0$，那么 $A^{\mathrm{T}}A$ 是半正定，由 6.5 节习题 22，rank $A^{\mathrm{T}}A$ = rank A .

11. 提示：将 A 的一个 SVD 写成形式 $A = U\Sigma V^{\mathrm{T}} = PQ$，其中 $P = U\Sigma U^{\mathrm{T}}$ 和 $Q = UV^{\mathrm{T}}$ 证明 P 是对称且有与 Σ 同样的特征值，解释为什么 Q 是正交矩阵.

13. a. 如果 $b = Ax$，那么 $x^+ = A^+b = A^+Ax$，由习题 12 (a)，x^+ 是 x 在 Row A 上的正交投影.

b. 从(a)和习题 12(c)，

$$Ax^+ = A(A^+Ax) = (AA^+A)x = Ax = b$$

c. 由于 x^+ 是 x 在 Row A 上的正交投影，由勾股定理

$$\|u\|^2 = \|x^+\|^2 + \|u - x^+\|^2$$

(c)立即得证.

15. [M] $A^+ = \frac{1}{40}\begin{bmatrix}-2 & -14 & 13 & 13\\ -2 & -14 & 13 & 13\\ -2 & 6 & -7 & -7\\ 2 & -6 & 7 & 7\\ 4 & -12 & -6 & -6\end{bmatrix}$，$\hat{x} = \begin{bmatrix}0.7\\ 0.7\\ -0.8\\ 0.8\\ 0.6\end{bmatrix}$

$\begin{bmatrix}A\\ x^{\mathrm{T}}\end{bmatrix}$的简化阶梯形除零以外的其他行外，与 A 的简化阶梯形一致，所以，将 A 中行乘上一个数量加到 x^{T} 可产生零向量，这说明 x^{T} 属于 Row A .

Nul A 的基：$\begin{bmatrix}-1\\ 1\\ 0\\ 0\\ 0\end{bmatrix}, \begin{bmatrix}0\\ 0\\ 1\\ 1\\ 0\end{bmatrix}$.

第 8 章

8.1

1. 参考答案：$y = 2v_1 - 1.5v_2 + 0.5v_3$，$y = 2v_1 - 2v_3 + v_4$，$y = 2v_1 + 3v_2 - 7v_3 + 3v_4$

3. $y = -3v_1 + 2v_2 + 2v_3$，权重之和为 1，故是一个仿射和.

5. a. 由于系数和为 1，$p_1 = 3b_1 - b_2 - b_3 \in \text{aff } S$

b. 由于系数和不为 1，$p_2 = 2b_1 + 0b_2 + b_3 \notin \text{aff } S$

c. 由于系数和为 1，$p_3 = -b_1 + 2b_2 + 0b_3 \in \text{aff } S$

7. a. $p_1 \in \text{Span } S$，但 $p_1 \notin \text{aff } S$.

b. $\boldsymbol{p}_2 \in \text{Span}\, S$,且 $\boldsymbol{p}_2 \in \text{aff}\, S$.

c. $\boldsymbol{p}_3 \notin \text{Span}\, S$,所以$\boldsymbol{p}_3 \notin \text{aff}\, S$.

9. $\boldsymbol{v}_1 = \begin{bmatrix} -3 \\ 0 \end{bmatrix}$ 且 $\boldsymbol{v}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$，可能有其他答案.

11. 见“学习指导”.

13. $\text{Span}\{\boldsymbol{v}_2 - \boldsymbol{v}_1, \boldsymbol{v}_3 - \boldsymbol{v}_1\}$ 是平面当且仅当 $\{\boldsymbol{v}_2 - \boldsymbol{v}_1, \boldsymbol{v}_3 - \boldsymbol{v}_1\}$ 线性无关. 设 c_2, c_3 满足 $c_2(\boldsymbol{v}_2 - \boldsymbol{v}_1) + c_3(\boldsymbol{v}_3 - \boldsymbol{v}_1) = 0$. 证明 $c_2 = c_3 = 0$.

15. 令 $S = \{\boldsymbol{x} : A\boldsymbol{x} = \boldsymbol{b}\}$，为证明 S 是仿射的，根据定理 3，只需证明 S 是一个平面. 令 $W = \{\boldsymbol{x} : A\boldsymbol{x} = \boldsymbol{0}\}$，根据 4.2 节的定理 2（或 2.8 节的定理 12），W 是 $\mathbb{R}^n$ 的一个子空间. 因为 $S = W + \boldsymbol{p}$，其中 $\boldsymbol{p}$ 满足 $A\boldsymbol{p} = \boldsymbol{b}$，根据 1.5 节的定理 6，$S$ 是 W 的平移，因此 S 是一个平面.

17. 一个恰当的集合包含任何三个向量，这三个向量不是共线的且 5 为它们的第三项. 若 5 是它们的第三项，则它们位于平面 $z = 5$. 若向量不是共线的，则它们的仿射包不可能是直线，故它一定是一个平面.

19. 若 $\boldsymbol{p}, \boldsymbol{q} \in f(S)$，则存在 $\boldsymbol{r}, \boldsymbol{s} \in S$ 使得 $f(\boldsymbol{r}) = \boldsymbol{p}$，$f(\boldsymbol{s}) = \boldsymbol{q}$. 对任意的 $t \in \mathbb{R}$，我们一定可以证明 $\boldsymbol{z} = (1-t)\boldsymbol{p} + t\boldsymbol{q}$ 在 $f(S)$ 中. 现在利用 $\boldsymbol{p}$ 和 $\boldsymbol{q}$ 的定义，及 f 是线性的这一事实. 完整的证明可参考“学习指导”.

21. 由于 B 是仿射的，定理 1 表明 B 包含 B 中的点的所有仿射组合，因此 B 包含 A 中的点的所有仿射组合，即 $\text{aff}\, A \subset B$.

23. 由于 $A \subset (A \cup B)$，由习题 22 知，$\text{aff}\, A \subset \text{aff}(A \cup B)$. 类似地，$\text{aff}\, B \subset \text{aff}(A \cup B)$，故 $[\text{aff}\, A \cup \text{aff}\, B] \subset \text{aff}(A \cup B)$.

25. 为证明 $D \subset E \cap F$，证明 $D \subset E$ 与 $D \subset F$. 完整的证明可参考“学习指导”.

8.2

1. 仿射相关且 $2\boldsymbol{v}_1 + \boldsymbol{v}_2 - 3\boldsymbol{v}_3 = \boldsymbol{0}$.

3. 集合是仿射无关的. 若点记为 $\boldsymbol{v}_1$，$\boldsymbol{v}_2$，$\boldsymbol{v}_3$，$\boldsymbol{v}_4$，则 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ 是 $\mathbb{R}^3$ 的基且 $\boldsymbol{v}_4 = 16\boldsymbol{v}_1 + 5\boldsymbol{v}_2 - 3\boldsymbol{v}_3$，但是线性组合中的系数和不为 1.

5. $-4\boldsymbol{v}_1 + 5\boldsymbol{v}_2 - 4\boldsymbol{v}_3 + 3\boldsymbol{v}_4 = \boldsymbol{0}$

7. 重心坐标是 $(-2, 4, -1)$.

9. 参考“学习指导”.

11. 当 5 个点的集合由减法转换，如去掉第一个点，4 个点的新集合一定是线性相关的，由 1.7 节定理 8，因为这 4 个点在 $\mathbb{R}^3$. 由定理 5，5 个点的原集合一定是仿射相关的.

13. 若 $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ 是仿射相关的，则存在不为零的 c_1, c_2，使得 $c_1 + c_2 = 0$ 且 $c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 = \boldsymbol{0}$. 证明这表明 $\boldsymbol{v}_1 = \boldsymbol{v}_2$. 相反地，假定 $\boldsymbol{v}_1 = \boldsymbol{v}_2$ 且选择特殊的 c_1, c_2 证明它们仿射相关. 细节参考“学习指导”.

15. a. 向量 $\boldsymbol{v}_2 - \boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ 且 $\boldsymbol{v}_3 - \boldsymbol{v}_1 = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$ 不是倍数关系，因此线性无关. 由定理 5，S 是仿射无关的.

b. $\boldsymbol{p}_1 \leftrightarrow \left(-\frac{6}{8}, \frac{9}{8}, \frac{5}{8}\right)$，$\boldsymbol{p}_2 \leftrightarrow \left(0, \frac{1}{2}, \frac{1}{2}\right)$，

$\boldsymbol{p}_3 \leftrightarrow \left(\frac{14}{8}, -\frac{5}{8}, -\frac{1}{8}\right)$，$\boldsymbol{p}_4 \leftrightarrow \left(\frac{6}{8}, -\frac{5}{8}, \frac{7}{8}\right)$，

$\boldsymbol{p}_5 \leftrightarrow \left(\frac{1}{4}, \frac{1}{8}, \frac{5}{8}\right)$

c. $\boldsymbol{p}_6$是$(-, -, +)$，$\boldsymbol{p}_7$是$(0, +, -)$，$\boldsymbol{p}_8$是$(+, +, -)$.

17. 假定 $S = \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_k\}$ 是仿射无关集合. 方程（7）有一个解，因为 $\boldsymbol{p}$ 在 $\text{aff}\, S$ 中. 因此方程（8）有一个解. 由定理 5，S 中点的齐次式是线性无关的，因此（8）有唯一解. 那么（7）也有唯一解，因为（8）包含两个方程，它们也在（7）中出现.

下面的讨论仿照 4.4 节定理 7 的证明. 如果 $S = \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_k\}$ 是仿射无关集合，那么由 $\text{aff}\, S$ 的定义，存在 $c_1, \cdots, c_k$ 满足（7）. 假设对于标量 $d_1, \cdots, d_k$，$\boldsymbol{x}$ 也可以表示为

$$\boldsymbol{x} = d_1\boldsymbol{b}_1 + \cdots + d_k\boldsymbol{b}_k,\ d_1 + \cdots + d_k = 1 \quad (7a)$$

那么相减得到：

$$\boldsymbol{0} = \boldsymbol{x} - \boldsymbol{x} = (c_1 - d_1)\boldsymbol{b}_1 + \cdots + (c_k - d_k)\boldsymbol{b}_k \quad (7b)$$

（7b）的系数和为 0，因为 c 和 d 的和分别为 1，这是可能的，除非（8）的每个系数都为 0，因为 S 是仿射无关集合.这就证明对于 $i=1,\cdots,k$，$c_i=d_i$.

19. 若 $\{\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3\}$ 是仿射相关集合,则存在不全为零的标量 c_1,c_2,c_3，满足 $c_1\boldsymbol{p}_1+c_2\boldsymbol{p}_2+c_3\boldsymbol{p}_3=\boldsymbol{0}$ 且 $c_1+c_2+c_3=0$. 再使用 f 的线性性质.

21. 设 $\boldsymbol{a}=\begin{bmatrix}a_1\\a_2\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}b_1\\b_2\end{bmatrix}$，$\boldsymbol{c}=\begin{bmatrix}c_1\\c_2\end{bmatrix}$，则

$$\det\begin{bmatrix}\tilde{\boldsymbol{a}} & \tilde{\boldsymbol{b}} & \tilde{\boldsymbol{c}}\end{bmatrix}=\det\begin{bmatrix}a_1 & b_1 & c_1\\a_2 & b_2 & c_2\\1 & 1 & 1\end{bmatrix}=\det\begin{bmatrix}a_1 & a_2 & 1\\b_1 & b_2 & 1\\c_1 & c_2 & 1\end{bmatrix},$$

由行列式的转置性质得到（3.2 节定理 5）. 由 3.3 节习题 30，行列式等于以 $\boldsymbol{a}$，$\boldsymbol{b}$，$\boldsymbol{c}$，为顶点的三角形面积的 2 倍.

23. 若 $\begin{bmatrix}\tilde{\boldsymbol{a}} & \tilde{\boldsymbol{b}} & \tilde{\boldsymbol{c}}\end{bmatrix}\begin{bmatrix}r\\s\\t\end{bmatrix}=\tilde{\boldsymbol{p}}$，由克莱姆法则，$r=\det\begin{bmatrix}\tilde{\boldsymbol{p}} & \tilde{\boldsymbol{b}} & \tilde{\boldsymbol{c}}\end{bmatrix}/\det\begin{bmatrix}\tilde{\boldsymbol{a}} & \tilde{\boldsymbol{b}} & \tilde{\boldsymbol{c}}\end{bmatrix}$. 由习题 21，这个商的分子是 $\triangle \boldsymbol{pbc}$ 面积的 2 倍，分母是 $\triangle \boldsymbol{abc}$ 面积的 2 倍. 由此证明了 r 的等式. 对于 s，t 的等式，也可使用克莱姆法则.

8.3

1. 见“学习指导”.

3. conv S 为空集.

5. $\boldsymbol{p}_1=-\frac{1}{6}\boldsymbol{v}_1+\frac{1}{3}\boldsymbol{v}_2+\frac{2}{3}\boldsymbol{v}_3+\frac{1}{6}\boldsymbol{v}_4=0$, 故 $\boldsymbol{p}_1\notin \text{conv } S$.

   $\boldsymbol{p}_2=\frac{1}{3}\boldsymbol{v}_1+\frac{1}{3}\boldsymbol{v}_2+\frac{1}{6}\boldsymbol{v}_3+\frac{1}{6}\boldsymbol{v}_4=0$, 故 $\boldsymbol{p}_2\in \text{conv } S$.

7. a. $\boldsymbol{p}_1$，$\boldsymbol{p}_2$，$\boldsymbol{p}_3$，$\boldsymbol{p}_4$ 的重心坐标分别为 $\left(\frac{1}{3},\frac{1}{6},\frac{1}{2}\right)$，$\left(0,\frac{1}{2},\frac{1}{2}\right)$，$\left(\frac{1}{2},-\frac{1}{4},\frac{3}{4}\right)$，$\left(\frac{1}{2},\frac{3}{4},-\frac{1}{4}\right)$.

   b. $\boldsymbol{p}_3$ 和 $\boldsymbol{p}_4$ 在 conv T 的外面. $\boldsymbol{p}_1$ 在 conv T 的内部. $\boldsymbol{p}_2$ 在 conv T 的边 $\overline{\boldsymbol{v}_2\boldsymbol{v}_3}$ 上.

9. $\boldsymbol{p}_1$ 和 $\boldsymbol{p}_3$ 在四面体 conv S 的外部. $\boldsymbol{p}_2$ 在包含顶点 $\boldsymbol{v}_2$，$\boldsymbol{v}_3$ 和 $\boldsymbol{v}_4$ 的面上. $\boldsymbol{p}_4$ 在 conv S 的内部. $\boldsymbol{p}_5$ 在 $\boldsymbol{v}_1$ 与 $\boldsymbol{v}_3$ 间的边沿上.

11. 见“学习指导”.

13. 若 $\boldsymbol{p},\boldsymbol{q}\in f(S)$，则存在 $\boldsymbol{r},\boldsymbol{s}\in S$，使得 $f(\boldsymbol{r})=\boldsymbol{p}, f(\boldsymbol{s})=\boldsymbol{q}$. 对 $0\leqslant t\leqslant 1$，证明线段 $\boldsymbol{y}=(1-t)\boldsymbol{p}+t\boldsymbol{q}$ 在 $f(S)$ 中. 现在利用 f 的线性及 S 的凸性来证明对某个 $\boldsymbol{w}\in S$ 有 $\boldsymbol{y}=f(\boldsymbol{w})$. 这也证明了 $\boldsymbol{y}\in f(S)$，即 $f(S)$ 是凸的.

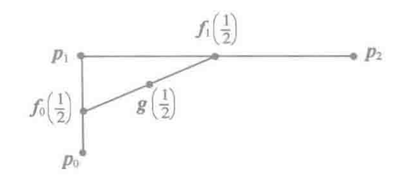
15. $\boldsymbol{p}=\frac{1}{6}\boldsymbol{v}_1+\frac{1}{2}\boldsymbol{v}_2+\frac{1}{3}\boldsymbol{v}_4$ 且 $\boldsymbol{p}=\frac{1}{2}\boldsymbol{v}_1+\frac{1}{6}\boldsymbol{v}_2+\frac{1}{3}\boldsymbol{v}_3$.

17. 假定 $A\subset B$，这里 B 是凸的，则由于 B 是凸的，定理 7 表明 B 包含了 B 中所有点的凸组合. 因此，B 包含 A 中所有点的凸组合，即 conv $A\subset B$.

19. a. 利用习题 18 证明 conv$A$ 与 conv$B$ 都是 conv$(A\cup B)$ 子集. 这将表明它们的组合也是 conv$(A\cup B)$ 的子集.

    b. 一种可能性是设 A 是正方形的两个相邻的角且 B 是另外两个角. 那么 $(\text{conv } A)\cup(\text{conv } B)$ 是什么，且 conv$(A\cup B)$ 又是什么呢？

21.



23. $$\begin{aligned}\boldsymbol{g}(t)&=(1-t)\boldsymbol{f}_0(t)+t\boldsymbol{f}_1(t)\\&=(1-t)\left[(1-t)\boldsymbol{p}_0+t\boldsymbol{p}_1\right]+t\left[(1-t)\boldsymbol{p}_1+t\boldsymbol{p}_2\right]\\&=(1-t)^2\boldsymbol{p}_0+2t(1-t)\boldsymbol{p}_1+t^2\boldsymbol{p}_2\end{aligned}$$

    $\boldsymbol{g}$ 的线性组合中权重之和为 $(1-t)^2+2t(1-t)+t^2$，它等于 $(1-2t+t^2)+(2t-2t^2)+t^2=1$. 当 $0\leqslant t\leqslant 1$ 时，权重在 0 与 1 之间，故 $\boldsymbol{g}(t)$ 在 conv$\{\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2\}$ 的内部.

8.4

1. $f(x_1,x_2)=3x_1+4x_2$ 且 $d=13$

3. a. 开的　　b. 闭的　　c. 两者都不是

   d. 闭的　　e. 闭的

5. a. 不是紧致的凸集
   b. 紧凸集
   c. 不是紧致的凸集
   d. 不是紧致的，也不是凸集
   e. 不是紧致的凸集

7. a. $\boldsymbol{n}=\begin{bmatrix}0\\2\\3\end{bmatrix}$或是其倍数
   b. $f(\boldsymbol{x})=2x_2+3x_3$且$d=11$

9. a. $\boldsymbol{n}=\begin{bmatrix}3\\-1\\2\\1\end{bmatrix}$或是其倍数
   b. $f(\boldsymbol{x})=3x_1-x_2+2x_3+x_4$且$d=5$

11. $\boldsymbol{v}_2$与$\boldsymbol{0}$在同一边，$\boldsymbol{v}_1$在另一边，$\boldsymbol{v}_3$在H中.

13. $\boldsymbol{p}=\begin{bmatrix}32\\-14\\0\\0\end{bmatrix}$，$\boldsymbol{v}_1=\begin{bmatrix}10\\-7\\1\\0\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}-4\\1\\0\\1\end{bmatrix}$.

15. $f(\boldsymbol{x})=x_1-3x_2+4x_3-2x_4$且$d=5$

17. $f(\boldsymbol{x})=x_1-2x_2+x_3, d=0$

19. $f(\boldsymbol{x})=-5x_1+3x_2+x_3, d=0$

21. 见“学习指导”.

23. $f(\boldsymbol{x})=3x_1-2x_2, 9<d<10$.

25. $f(x,y)=4x+1, d=12.75$,即$f(3,0.75)$. 点(3,0.75)是$B(\boldsymbol{0},3)$中心到$B(\boldsymbol{P},1)$中心距离的$\frac{3}{4}$.

27. 8.3 节习题 2（a）给出一种可能性. 或令$S=\{(x,y):x^2y^2=1且y>0\}$，则 conv S 是上方的（开）半平面.

29. 令$\boldsymbol{x},\boldsymbol{y}\in B(\boldsymbol{p},\delta)$，假设$\boldsymbol{z}=(1-t)\boldsymbol{x}+t\boldsymbol{y}$，其中$0\leqslant t\leqslant 1$，则

$$\begin{aligned}\|\boldsymbol{z}-\boldsymbol{p}\|&=\|[(1-t)\boldsymbol{x}+t\boldsymbol{y}]-\boldsymbol{p}\|\\&=\|(1-t)(\boldsymbol{x}-\boldsymbol{p})+t(\boldsymbol{y}-\boldsymbol{p})\|<\delta\end{aligned}$$

8.5

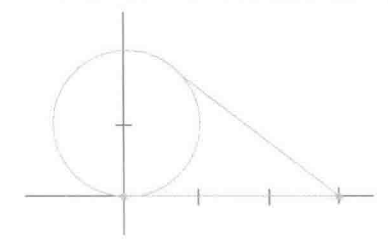
1. a. 在点$\boldsymbol{p}_1, m=1$　　b. 在点$\boldsymbol{p}_2, m=5$
   c. 在点$\boldsymbol{p}_3, m=5$

3. a. 在点$\boldsymbol{p}_3, m=-3$
   b. 在集合conv$\{\boldsymbol{p}_1,\boldsymbol{p}_3\}$，$m=1$
   c. 在集合 conv$\{\boldsymbol{p}_1,\boldsymbol{p}_2\}$，$m=-3$

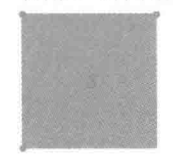
5. $\left\{\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}5\\0\end{bmatrix},\begin{bmatrix}4\\3\end{bmatrix},\begin{bmatrix}0\\5\end{bmatrix}\right\}$

7. $\left\{\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}7\\0\end{bmatrix},\begin{bmatrix}6\\4\end{bmatrix},\begin{bmatrix}0\\6\end{bmatrix}\right\}$

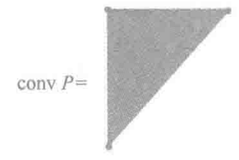
9. 原点是极端点，但不是顶点. 解释原因.



11. 一种可能是令S是一个正方形，包括边界但不包含所有的边界. 例如，只包括两条相邻边. 轮廓P的凸包是一个三角形区域.



conv P=



13. a. $f_0(C^5)=32, f_1(C^5)=80, f_2(C^5)=80, f_3(C^5)=40, f_4(C^5)=10$，且$32-80+80-40+10=2$.

   b.

	f_0	f_1	f_2	f_3	f_4
S^1	2				
S^2	4	4			
S^3	8	12	6		
S^4	16	32	24	8	
S^5	32	80	80	40	10

   对于一般的公式，见“学习指导”.

15. a. $f_0(P^n)=f_0(Q)+1$
   b. $f_k(P^n)=f_k(Q)+f_{k-1}(Q)$
   c. $f_{n-1}(P^n)=f_{n-2}(Q)+1$

17. 见“学习指导”.

19. 设 S 是凸集且令 $x \in cS + dS$，这里 $c>0, d>0$，则存在 S 中的 s_1, s_2，满足 $x = cs_1 + ds_2$. 但是

$$x = cs_1 + ds_2 = (c+d)(\frac{c}{c+d}s_1 + \frac{d}{c+d}s_2)$$

现在证明等式右边的表达式是 $(c+d)S$ 中的元素.

相反地，在 $(c+d)S$ 中取一点，证明该点在 $cS+dS$ 中.

21. 提示：假定 A,B 是凸集. 设 $x, y \in A+B$，则存在 $a,c \in A$，且 $b,d \in B$，满足 $x = a+b$ 且 $y = c+d$. 对任意 t 满足 $0 \leqslant t \leqslant 1$，证明

$$w = (1-t)x + ty = (1-t)(a+b) + t(c+d)$$

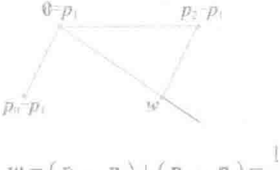
代表 $A+B$ 中的一点.

8.6

1. $x(t)+b$ 的控制点为 p_0+b，p_1+b，p_3+b. 写出过这些点的贝塞尔曲线，且用代数法证明曲线是 $x(t)+b$. 见“学习指导”.

3. a. $x'(t) = (-3+6t-3t^2)p_0 + (3-12t+9t^2)p_1 + (6t-9t^2)p_2 + 3t^2p_2$，故 $x'(0) = -3p_0 + 3p_1 = 3(p_1 - p_0)$ 且 $x'(1) = -3p_2 + 3p_3 = 3(p_3 - p_2)$. 这证明切向量 $x'(0)$ 由 p_0 指向 p_1 且是 $p_1 - p_0$ 长度的 3 倍. 类似地，$x'(1)$ 由 p_2 指向 p_3 且是 $p_3 - p_2$ 长度的 3 倍. 尤其 $x'(1) = 0$ 当且仅当 $p_3 = p_2$.

   b $x''(t) = (6-6t)p_0 + (-12+18t)p_1 + (6-18t)p_2 + 6tp_3$，$x''(0) = 6p_0 - 12p_1 + 6p_2 = 6(p_0 - p_1) + 6(p_2 - p_1)$ 且 $x''(1) = 6p_1 - 12p_2 + 6p_3 = 6(p_1 - p_2) + 6(p_3 - p_2)$. 对 $x''(0)$ 的图像，构造以 p_1 为原点的坐标系，暂且把 p_0 记为 $p_0 - p_1$，把 p_2 记为 $p_2 - p_1$. 最终，构造由新的原点到 $p_0 - p_1$ 与 $p_2 - p_1$ 的和的点的直线. 此直线指向 $x''(0)$.



$$w = (p_0 - p_1) + (p_2 - p_1) = \frac{1}{6}x''(0)$$

5. a. 由习题 3 或方程 (9)

$$x'(1) = 3(p_3 - p_2)$$

使用 $x'(0)$ 的方程，来自 $y(t)$ 的控制点，且得 $y'(0) = 3p_3 + 3p_4 = 3(p_4 - p_3)$

由 C^1 的连续性，$3(p_3 - p_2) = 3(p_4 - p_3)$，故 $p_3 = (p_4 + p_2)/2$ 且 p_3 是由 p_2 到 p_4 的线段的中点.

   b. 若 $x'(1) = y'(0) = \mathbf{0}$，则 $p_2 = p_3$ 且 $p_3 = p_4$. 因此，由 p_2 到 p_4 的线段就是点 p_3.（注意：在此情形下，由定义知，组合曲线仍是 C^1 连续的，然而其他控制点 p_0，p_1，p_5，p_6 的选取也可以产生在 p_3 具有可视角的曲线，此情形下曲线在 p_3 不是 G^1 连续的.

7. 提示：使用习题 3 的 $x''(t)$ 且把第二条曲线视为：

$$y''(t) = 6(1-t)p_3 + 6(-2+3t)p_4 + 6(1-3t)p_5 + 6tp_6$$

则设 $x''(1) = y''(0)$. 由于曲线在 p_3 是 C^1 连续的，习题 5 (a) 证明了 p_3 是由 p_2 到 p_4 的线段的中点. 这表明 $p_4 - p_3 = p_3 - p_2$. 使用此置换来证明 p_4 与 p_5 由 p_1，p_2，p_3 唯一确定. 只有 p_6 是可任意选取的.

9. 写出多项式 $x(t)$ 的系数的向量形式，扩展此多项式的系数，且把向量记为 $M_B u(t)$：

$$\begin{bmatrix} 1-4t+6t^2-4t^3+t^4 \\ 4t-12t^2+12t^3-4t^4 \\ 6t^2-12t^3+6t^4 \\ 4t^3-4t^4 \\ t^4 \end{bmatrix},$$

$$= \begin{bmatrix} 1 & -4 & 6 & -4 & 1 \\ 0 & 4 & -12 & 12 & -4 \\ 0 & 0 & 12 & -12 & 6 \\ 0 & 0 & 4 & 4 & -4 \\ 0 & 0 & 0 &) & 1 \end{bmatrix} \begin{bmatrix} 1 \\ t \\ t^2 \\ t^3 \\ t^4 \end{bmatrix}$$

$$M_B = \begin{bmatrix} 1 & -4 & 6 & -4 & 1 \\ 0 & 4 & -12 & 12 & -4 \\ 0 & 0 & 6 & -12 & 6 \\ 0 & 0 & 0 & 4 & -4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

11. 见“学习指导”.

13. a. 提示：使用 $q_0 = p_0$.

b. 把方程（13）的第一与最后一部分乘以$\frac{8}{3}$以解$8\boldsymbol{q}_2$.

c. 使用方程（8）以替换$8\boldsymbol{q}_3$且应用（a）.

15. a. 由方程（11），$\boldsymbol{y}'(1)=0.5\boldsymbol{x}'(0.5)=\boldsymbol{z}'(0)$.

b. 观察$\boldsymbol{y}'(1)=3(\boldsymbol{q}_3-\boldsymbol{q}_2)$. 由方程（9），用$\boldsymbol{y}(t)$及它的控制点替代$\boldsymbol{x}(t)$及它的控制点.类似地，对于$\boldsymbol{z}(t)$及它的控制点，$\boldsymbol{z}'(0)=3(\boldsymbol{r}_1-\boldsymbol{r}_0)$. 由（a），$3(\boldsymbol{q}_3-\boldsymbol{q}_2)=3(\boldsymbol{r}_1-\boldsymbol{r}_0)$. 用$\boldsymbol{q}_3$代替$\boldsymbol{r}_0$，则得$\boldsymbol{q}_3-\boldsymbol{q}_2=\boldsymbol{r}_1-\boldsymbol{q}_3$，因此$\boldsymbol{q}_3=(\boldsymbol{q}_2+\boldsymbol{r}_1)/2$.

c. 设$\boldsymbol{q}_0=\boldsymbol{p}_0$且$\boldsymbol{r}_3=\boldsymbol{p}_3$. 计算$\boldsymbol{q}_1=(\boldsymbol{p}_0+\boldsymbol{p}_1)/2$. 计算$\boldsymbol{r}_2=(\boldsymbol{p}_2+\boldsymbol{p}_3)/2$. 计算$\boldsymbol{m}=(\boldsymbol{p}_1+\boldsymbol{p}_2)/2$. 计算$\boldsymbol{q}_2=(\boldsymbol{q}_1+\boldsymbol{m})/2$，$\boldsymbol{r}_1=(\boldsymbol{m}+\boldsymbol{r}_2)/2$. 计算$\boldsymbol{q}_3=(\boldsymbol{q}_2+\boldsymbol{r}_1)/2$且设$\boldsymbol{r}_0=\boldsymbol{q}_3$.

17. a. $\boldsymbol{r}_0=\boldsymbol{p}_0$，$\boldsymbol{r}_1=\dfrac{\boldsymbol{p}_0+2\boldsymbol{p}_1}{3}$，$\boldsymbol{r}_2=\dfrac{2\boldsymbol{p}_1+\boldsymbol{p}_2}{3}$，$\boldsymbol{r}_3=\boldsymbol{p}_2$.

b. 提示：写出本节（7）的标准公式，且对于$i=0,\cdots,3$，用$\boldsymbol{r}_i$替代$\boldsymbol{p}_i$，然后分别用$\boldsymbol{p}_0$，$\boldsymbol{p}_2$替代$\boldsymbol{r}_0$，$\boldsymbol{r}_3$：

$$\begin{aligned}\boldsymbol{x}(t)&=(1-3t+3t^2-t^3)\boldsymbol{p}_0\\&\quad+(3t-6t^2+3t^3)\boldsymbol{r}_1\\&\quad+(3t^2-3t^3)\boldsymbol{r}_2+t^3\boldsymbol{p}_2\end{aligned}\qquad(iii)$$

使用（a）中$\boldsymbol{r}_1$，$\boldsymbol{r}_2$的公式以检验$\boldsymbol{x}(t)$的表达式中的第二项与第三项.

线性代数及其应用（原书第4版）

本书根据“线性代数课程研究小组”的建议，通过认真观察学生的实际需要和许多不同专业使用线性代数知识的共同点而选材。

本书是一本优秀的现代教材，给出线性代数最新的基本介绍和一些有趣应用，目的是帮助学生掌握线性代数的基本概念及应用技巧，为后续课程的学习和工作实践奠定基础。第4版新增了一章“向量空间的几何学”，在第3版的基础上为学生和教师提供了进一步的技术支持。

本书特点

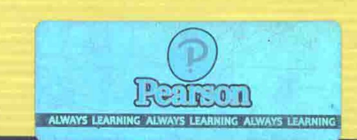
- 介绍了线性代数的基本概念、理论和证明，包含大量例题、练习题、习题等，广泛选取的应用说明了线性代数的作用，可以用于在工程学、计算机科学、物理学、数学、生物学、经济学和统计学中解释基本原理和简化计算。
- 提前介绍重要概念，许多基本概念含在每章开始的“介绍性实例”中，然后从不同的观点逐步深入讨论。
- 矩阵乘法采用了现代观点，本书在定义和证明中处理的是矩阵的列，而不是矩阵的元素，这种现代方法简化了许多论据，且将向量空间思想和线性系统的研究联系在一起。
- 结合应用数学软件，强调了计算机对科学和工程学中线性代数的发展和实践的影响。“数值计算的注解”指出了数值计算中出现的问题，以及理论概念（如矩阵求逆）和计算机实现（如LU分解）之间的区别。
- 可以从网站www.pearsonhighered/lay上找到相关的技术支持。

作者简介

戴维 C. 雷（David C. Lay）在美国加利福尼亚大学获得硕士和博士学位。他是马里兰大学帕科学院数学系教授，同时还是阿姆斯特丹大学、阿姆斯特丹自由大学和德国凯泽斯劳滕大学的访问教授。Lay教授是“线性代数课程研究小组”的核心成员，发表了30多篇关于泛函分析和线性代数方面的论文，并与他人合著多部数学教材。

Linear Algebra and Its Applications

Fourth Edition



上架指导：数学

ISBN 978-7-111-55978-8

www.pearson.com

定价：79.00元

投稿热线：(010) 88379604
客服热线：(010) 88378991 88361066
购书热线：(010) 68326294 88379649 68995259
封面设计·杨宇梅

华章网站：www.hzbook.com
网上购书：www.china-pub.com
数字阅读：www.hzmedia.com.cn